الإحصاء للمكتبيين



تأليف

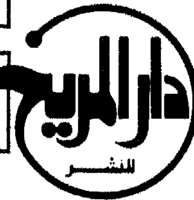
راي ل. كاربنتر
جامعة شمال كارولينا
شابل هيل

إلين ستوري فاسو
جامعة شمال كارولينا
شابل هيل

ترجمة

الدكتور/ سيّد حسب الله
قسم علوم المكتبات والمعلومات
جامعة الملك سعود - الرياض

الدكتور/ محمد جلال سيد محمد غندور
قسم المكتبات والمعلومات
جامعة القاهرة - بني سويف



ص. ب: ١٠٧٢٠ - الرياض: ١١٤٤٣ - فاكس: ٤٦٥٧٩٣٩
المملكة العربية السعودية - تلفون ٤٦٥٨٥٢٣ - ٤٦٤٧٥٣١

الإحصاء للمكتبيين

ردمك : ١ - ٤٥٧ - ٢٤ - ٩٩٦٠


بسم الله الرحمن الرحيم

المحتويـــات

جــــــداول:

المقدمة

تمثل عملية البحث العلمي، رؤية فكرية فعالة ومتميزة، فهي تساعد على فهم وإدراك كيفية تفاعل الأحداث التي تجمعها علاقات مشتركة، وفي كثير من الأحيان تمدنا بالسبل التي تعيننا على التكيف مع البيئة المحيطة بنا، وتشتمل عملية البحث العلمي على كل مناهج جمع البيانات الرقمية وطرق الوصول إلى استنتاجات تتعلق بخصائص وسلوكيات هذه بالأرقام ويتحدد نمط البيانات المختارة لدراسة ما، بالتطورات السابقة التي حدثت في المجال موضع الدراسة، كما يتضمن علم الإحصاء المناهج المستخدمة لتنظيم عملية البحث للبيانات الرقمية، وكيفية استغلالها للوصول إلى نتائج فعالة.

وتعد معرفة كيفية إجراء البحوث، على درجة كبيرة من الأهمية للمكتبيين، وخبراء الأرشيف وأخصائي المعلومات أكثر من غيرهم من المتخصصين، وذلك لاعتبارات كثيرة. ومن الضروري أن نتعرف على ماهية البحث العلمي، ودور المكتبيين فيه، وذلك حتى نتفهم هذه الحقيقة (لأغراض هذا البحث، ولتجنب تكرار المصطلحات، سترد كلمة «مكتبيين» لتشمل علماء المعلومات، وخبراء الأرشيف، الوثائقيين).

أدت التطورات الكبيرة التي حدثت في العالم، والتعقيدات التي نشأت عن تداخل المفاهيم، وتشابك الأنظمة، إلى قيام أنماط من المؤسسات ذات أهمية خاصة، لنا نحن المكتبيين العاملين في مجال المعلومات إلى ثلاثة أضعاف ماهي عليه، وأصبح دور المكتبيين كوسطاء مابين المعلومات المسجلة ومستخدميها، ليس قاصراً فقط على إيجاد وتحديد مكان المعلومات، بل يتعدى ذلك إلى تحليل تلك المعلومات وتقييمها وتفسيرها لصالح المستفيد. لأن معظم المعلومات ترد في أوعية تأخذ شكل البحوث والمقالات والتقارير التي يتحتم على المكتبيين التعرف عليها واختبارها حتى يتمكنوا من بثها، ويستطيع المكتبي أن يقوم بهذا الدور بكفاءة عالية، إذا ما تمكن من إجادة اللغات،

والتعرف على المفاهيم الأساسية، وتفهم المناهج البحثية، التي صيغ بها هذا الإنتاج الفكري. إن معرفة هذه الإجراءات البحثية تساعد المكتبي كثيراً على تقديم خدمات معلومـات متـوازنة، وأكثر فائدة، وأكثر جدية، وبذلك يستطيع أن يصل ـ وبدرجة كبيرة ـ إلى تحقيق أحد الأهداف الرئيسية للتخصص.

إضافة إلى ذلك، يوجد بُعدين آخرين لدور المكتبي، أولهما: أن المكتبي ـ نفسه ـ يُعـد مستخـدمـا لشتي أنواع البيانات، ويستفيد من دراستها في تطوير ورفع كفاءته المهنية، وذلك لتقديم خدمات معلومات أفضل للمؤسسة التي ينتمي إليها (وهو الأمر الـذي سيناقش بعد قليل)؛ وثانيهما: أنه بالرغم من أن عدد المكتبيين المشاركين في مشاريع الأبحاث العلمية في الوقت الحاضر يعتبر متواضعاً إلا أن هذا العدد يتوقع له الزيادة، بدرجة محسوسة في القريب العاجل، وقد تكون هذه المشاريع البحثية متعلقة بالأنشطة الرئيسية لمجال المكتبات والمعلومات.

إن المكتبيين قد يجدوا أنفسهم مشـاركـين في أبحاث ودراسات تتعلق بمجالات أخرى، فليس بمستبعد مثلاً أن يجد المكتبيون القائمون على أمر المكتبات المدرسية والعامة ـ أنفسهم مساهمين في دراسات تتعلق بالتعليم في مجتمعهم المحلي، مما يستدعي تعـاملهم مع مجالات معلومـات واسعـة النـطاق، تتعلق بالعـوامـل الاجتـماعية، والاقتصادية، والسياسية، والتكنولوجية لمجتمعاتهم. وقد يجد المكتبيون في المكتبات الأكاديمية والمتخصصة، أنفسهم في ظروف مشابهة، تحتم عليه المشاركة في المجالات الـدراسية والأنشـطة البحثية للمؤسسـات التي ينتمون إليها، في إطار التخصصات الموضوعية لمؤسساتهم.

سيجد المكتبي ـ كمستخدم للمعلومات ـ أن البيانات الإحصائية، أصبحت جزءاً هامـاً من الإنتـاج الفكـري للمجـالات التي يتعـامل معها، فالكتب والدوريات، والأشكال الأخرى من مصادر المعلومات للمجالات المعرفية المختلفة، كإدارة الأعمال وعلم الاجتماع، أو العلوم السياسية، أصبحت أقرب ما تكون إلى الدراسات الكمية. وفي واقـع الأمـر، يتحتم على المكتبيين معـرفـة كيفية التعامل مع البيانات الرقمية، كبـاحثين نشطين، سواء بصفتهم مصدراً للمعلومات، أو كمحققين وفاحصين لها. لذلك يُعني هذا الكتاب بدراسة بعض الإجراءات الخاصة بتحليل البيانات التي تُعرف بإحصاءات العلوم الاجتماعية، وكما ذكر سابقاً، فإن المكتبيين يقضون جل وقتهم في دراسـة أبحـاث الإنتاج الفكري في مجال علوم المكتبات، والمجالات الأخرى ذات

العـلاقـة ويُتوقع منهم دائماً أن يبذلوا أقصى جهد في مجالات البحوث التي تهتم بها مؤسساتهم، وأحد الأسباب المهمة التي أدت إلى ذلك، هو ازدياد أهمية البحوث التي تجري في مجال الإدارة، وكذلك ازدياد أهمية الإدارة في مساعدة المكتبي للقيام بدوره على أكمل وجه.

وإذا كان البحث يقـود إلى التـوصل لمعارف جديدة، فإن تلك المعارف ذات أهمية قصـوى للقيام بعمليات تحليل واسعـة النـطاق، سواء في مجال المكتبـات أو شبكـات المعلومـات ومـراكـز المعلومـات، كوحـدات معلومـات، وهـذا يعني أن الاعتماد على المعلومـات، يُعـد القاعدة الأساسية للإدارة العلمية الرشيدة. إن هدف الأبحاث هو تزويد المكتبيين والمستفيدين من المكتبـات بالمعلومات الدقيقة الموثوق بها، لتساعدهم على اتخاذ القرار الأفضل، وتحدد جودة البيانات (المعلومات) المستخرجة من الأبحاث، طبقـاً لأهميتهـا، ومـدى مشـاركتهـا في زيادة فعـالية الخـدمات المكتبية، ورفع كفاءة المستفيدين من المكتبـات. وفي هذا الصدد، يؤمل رفع جودة إدارة المعلومات وكفاءتها عن طريق المعارف المستمدة من التحليلات الإحصائية. وقد أفرزت الظروف السياسية والاجتـماعية والاقتصـادية في الآونة الأخيرة، مبدأ مهماً سميّ «بالمسؤولية العامة»، كما تزايدت في العقـدين الأخـيرين، الحاجة إلى المكتبات وتدعيمها، للدرجة التي دفعت المؤسسـات الإنتـاجية الممـولة في المجتمع للدعوة إلى تطوير ودعم الخدمات المكتبية. وكنتيجـة لتعـاظم دور المكتبات في تدعيم القطاع الأقتصادي، تم تبني شعار «دعهم يفعلون Laisser Faire[1]» في مجال إدارة المكتبات، للتعبير عن الدعم والمساندة التي تجدها المكتبـات للقيام بدور فعـال في المجالات الإنتاجية بالمجتمع، ولا ندعي هنا أن مبدأ «المسؤولية العـامـة» خاص بقـطاع المكتبـات ـ فقط ـ بل هو مبدأ ينسحب على جميع المؤسسات في المجتمع. نجد في الوقت نفسه أن حجم المشكلات الاقتصادية ونوعيتها التي واجهت المكتبات، أخذت في التزايد، وظهرت مشكلات لم تكن معروفة من قبل في هذا المجال، مشكلات تتعلق بارتفاع تكلفة أوعية المعلومات وارتفاع تكلفة الكوادر المؤهلة. كما برزت مشكـــلات أخـرى تتعلق بأدوات الضبط الببليوجـرافي، وتـداول المعلومات، وهي نماذج لأمثلة عديدة كان للبحث العلمي دور كبير في إيجاد الحلول لها، ليس فقط على صعيد مبدأ «المسؤولية العامة»، بل ـ أيضاً ـ للمساعدة في تحقيق الهدف المهني الرئيسي للمكتبات، وهو «التحسين المستمر في الخدمات المكتبية، لمقابلة التغير

(1) وهو مصطلح فرنسي يعني: أن يترك لهم الأمر ليفعلوا ما يشاؤون [المترجـان].

الدائم في طبيعة مصادر المعلومات وحاجات المستفيدين» . ولكن كل المبررات والأسباب التي تساق لإبراز مزايا المكتبات لن تلقي أذاناً صاغية حتى من مناصري ومؤيدي تدعيم خدمات المكتبات، إذا لم تستند هذه المبررات على بحوث ودراسات منطقية، تعتمد في الأساس على التحليلات الكيفية والكمية.

وفي الحقيقـة، هناك عاملان أساسيان يساعدان على زيادة قدرة المكتبات في القيام بدورهـا في إجراء البحوث وتقييمها : العامل الأول يتعلق بالمشكلات التي تنشأ نتيجة للتطورات السريعة التي تحدث في مجال تقنيات الاتصالات والتكنولوجيا . والعامل الثاني عن التوقعات البحثية المتعلقة، بالمسؤولية العامة والمهنية المتخصصة، التي تأخذ شكل البيانـات الإحصائية . وهناك دلائل قوية، تشير إلى التزايد المستمر في أعداد المكتبيين المؤهلين القادرين على فهم عالمهم المهني، واستيعاب دورهم في حل مشاكل مجتمعاتهم، هـؤلاء المكتبيون استحدثوا طرقاً وأساليبا إحصائية قاموا بتطبيقها بنجاح على المشكلات التي تسببت فيها الظروف الاجتماعية والتكنولوجية التي تمر بها مجتمعاتهم، وقد ساعدوا في إيجاد حلول لمشكلات واجهت المؤسسات الأم التي تخدمها تلك المكتبات.

تُعـد الأبحـاث ـ التي تتميز بقـدر معقـول من الجـودة ـ قاعـدة صلبـة للدراسات التقـويمية وهـذه المسـألـة تستحق منـا البحث والتمحيص؛ فقد توصلت العديد من التجارب بأن نتائج الأبحاث لا يعتد بها إلا إذا كانت قائمة على مفاهيم تنظيرية تنطلق ـ من حيث المبـدأ ـ من أسـاس نظري . استناداً على هذا المفهوم، يتشكل البحث في البداية من كيان نظري، أو على الأقل، من أحد أشكال المفاهيم المجردة، ثم يتحول إلى دراسة تجريبية، تضيء لنا الطريق، في صورة برهان قاطع، أو وصف دقيق للمفاهيم التنـظيرية المطروحة، ويمكن صياغة هذه المسلّمة، بطريقة أفضل على النحو التالي : يتحتم اختبار المفاهيم المجردة عن طريق المناهج التجريبية، وليتحقق ذلك، يجب اتباع التفكـير العلمي القـائم على الحجج البراهين العلمية، والاختيار الدقيق للإجراءات البحثية المناسبة.

هناك رؤية أخرى، قائمة على دراسة حقائق عالم اليوم، تشير إلى أنه ليس بالضرورة أن يستند البحث العلمي على مفاهيم تنظيرية مسبقة . وقد أوردنا سابقاً، أن الإجراءات البحثية تُعد وسيلة متميزة للتعايش، والانخراط، والتأقلم مع البيئة المحيطة . ولذا فقد تتخذ إجراءات البحوث العلمية مساراً يبدأ من صيغ ومفاهيم تجريدة، أو يستند على مجرد حس علمي بسيط، أو بالإثنين معاً . حقاً أن قيمة تقنيات البحث ونتائجها لا تنبع

فقط من كونها أداة تسمح لنـا بدراسـة متأنية ودقيقة للحقائق، أو لتأكيد مسلمات أو مفاهيم مجردة؛ بل أن الإجراءات البحثية تُعد ـ في حد ذاتها ـ قيمة كبيرة. وإذا ما اقتنعنا وآمنـا بالنتـائـج والمسلمات العلمية، التي يتـوصـل إليهـا العلماء والباحثون المتميزون بالعقلانية، والنشاط، والفضول العلمي الجاد فإن إجراءات بحوثهم تمكننا من تجسيد مفاهيمنا، وتطور من تفكيرنا العقلاني، وتفسر، وتبرر لنا الظواهر بطريقة منطقية. وفي هذا المجـال تسـاعـدنـا تقنيات البحث والتحليل الإحصـائى، على وضع مفاهيمنا، ومفاهيم العالم من حولنا في الأطر المناسبة لها. وهذا يشمل أعمال المكتبات وخدماتها، التي لا يمكن تحقيقها بالمفاهيم المجردة فقط.

يُعد البحث عن الحقيقة والصدق شيئان أساسيان في مجال البحوث. ونقاط البداية للانطلاق نحو هذا البحث، متعددة. أولها ضرورة الاعتراف بالحاجة لاختيار مفاهيمنا عن طريق إجـراء الـدراسـات، وبـالتـالي حاجتنـا إلى وضـع التصـور السليم لهذه الدراسات، ذلك التصور المستند على تحديد الدراسات وتأطيرها منذ البداية بالمفاهيم العقـلانية. وثـانيهـا: تقـوم على التسليم بأن البحث يمكن ان يتبـع بالحس العلمى الـلامـادى، بشرط أن يتم تشكيل الإطـار والإجـراءات البحثية على أساس المفاهيم العلمية الـواضحـة. ومن الأهمية بمكـان التفـرقـة، بين هذا الرأي وذاك، فالتشوش الموضوعي (المتمثل في الإفراط في استخدام الحس العلمي) يقود إلى الإفراط في النواحي الشكلية الخاصة بالبحث (وليس من المستحب ألا يكون هناك أساس نظري للبحث)؛ من ناحية أخـرى، لا يجب إعـاقـة البحث القـائم على الحس العلمي، بالإكثـار من المعـوقـات المتمثلة في الاستغـراق في استخدام البنائيات المنهجية (فهذا أيضا يقود إلى إبحاث غير متقنة، وغير ذات قيمة). لذا، فإنه يمكن، بل ويجب، أن تجرى البحوث ـ بقـدر المستـطـاع ـ في إطار مفاهيم ونظريات محددة ودقيقة. ومن ناحية أخرى، فإن البحوث المستندة على الحد الأدنى من المفاهيم التنظيمية يمكن أن ينتج عنها معلومات مفيدة، وهذه بدورها تقود إلى إعادة تعريف وتوضيح وتفهم أفضل للمفاهيم.

يفتقد علم المكتبات ـ حتى الآن ـ الإطار النظري التنظيمي المتطور، الذي يساعد على تفسير أغراض وأنشطة هذا المجال التخصصي وشرحها. ويعمل الكثير من المكتبيين والباحثين والعلماء على تطوير النظريات والقوانين المتعلقة بسلوكيات تخصص المكتبات والمعلومات، إضافة إلى ذلك فـإن فـي الكم الكبير مـن النظريات والمفاهيم التي تتعلق بالدراسات في مجال السلوكيات الإنسانية، وعلم الإدارة، وعلم المعلومات، يوجد نبع هائل من الأفكار المبتكرة والجديدة، التي لا تحتاج إلا إلى بعض التعديلات والإضافات

البسيطة، لتستخدم، بعد ذلك، بنجاح في مجال البحث العلمي للمكتبات، بغرض تطوير أساسه الفكري والتنظيري، فيستطيع المكتبيون ـ في هذا المجال ـ الاستفادة من النمـو المـطرد في الدراسات المتعلقة بالمؤسسات المناظرة، وخاصة تلك التي تنتمي إلى قطاع الخدمات العامة، حيث يقومون بفحص فروض نتائج تلك الدراسات ومعطياتها وتطويعها لتناسب السمات الخاصة بالمكتبات ومراكز المعلومات. إن فهم وتطبيق نتائج البحـوث ذات الصلة بالمجـالات المعـرفية المختلفة، إضافة إلى تلك التي تخص مجال المكتبات والمعلومات، يؤدي إلى وجود مخزون كبير من الأفكار القابلة للبحث، تكفي (إن لم تكن تزيد) لتغطية جميع الجوانب البحثية للتخصص.

نأمـل أن نُأصّـل مفهـومـاً واضحاً لعلم الإحصاء، والتحليل الإحصائي، ومدى فائدتها لتخصص المكتبات والمعلومات في الفصول التالية من هذا الكتاب. وقد راعينا خلال استعـراضنـا لهذه المـوضوعات أن نتعامل مع الحد الأدنى المطلوب للمعطيات الـرياضية، فليس من أهـداف بحثنـا هذا، أن نؤهـل القـراء ليصبحـوا احصـائيين متخصصين، وهو الأمر الذي لا نستطيع تحقيقه من خلال هذا العمل. لذا، فهدفنا هو إحـاطـة القـراء علماً بالتقنيات الإحصـائية الأكثـر شيوعـاً واستخـدامـاً، وتعـريفهم بالمصـطلحـات الرئيسية المستخدمة في أي من مجالات البحث الرئيسية، ويتعين على القارىء، بعد مطالعته لهذا العمل، وتصديه لحل المعضلات التي أوردناها في آخر كل فصـــل فيه، أن يكتسب المعلومـات الأسـاسية المتعلقـة بالمبـادىء الـرئيسية للتحليل الإحصـــائي، أمـــا هؤلاء الـــذين يودون الاستـزادة والتعمق حول علم الإحصـــاء، والموضوعات المنهجية المتصلة به، فقد أوردنا قائمة مراجع في نهاية كل فصل، لتعينه على هذا الأمر.

تبدأ مناقشاتنا في الفصل الأول بمقدمة عن الإحصاء الوصفي، التي سرعان ما تتجه نحو تحديد الخصائص المميزة لمجموعة أشخاص، أو مجموعة حالات بحث، على ضوء متغـيرات تم اختيارها لتناسب أغراض البحث. ومن أجل إحصاء هذه الخصائص، فقد تعرضنا أيضاً في الفصل الأول إلى إجراءات القياس، التي تعتمد في الأساس على إحصاء وتصنيف الأفراد بناء على ما يحوزنه من متغيرات أو ما يتصفون به من خصائص.

يحتـوي الفصل الأول ـ أيضاً ـ على مناهج تعالج كيفية حساب بعض الملخصات الإحصـائية، ومن خلال هذه المنـاهـج، نستطيع أن نلخص كم هائل من المعلومات باستخـدام كم متـواضـع من الإحصاءات والأرقام. ومع ذلك، فاهتمامنا ـ في بعض

الأحيان ـ لا يتركز في الخروج بنتائج تحليل تختص بعدد محدود من الأفراد أو من حالات البحث التي نملك عنها معلومات كافية، بل نفضل تعميم النتائج لتشمل على قاعدة أكبر من أفراد مجتمع البحث (مثل جميع المستفيدين من المكتبات ومراكز المعلومات)، وفي مثل هذه الحالة، نبني استنتاجنا على النتائج التي توصلنا إليها من تحليل البيانات المتوافرة لدينا عن عينة مجتمع البحث موضع الدراسة، ولكن لإعطاء هذه النتائج صفة الدقة والمصداقية، يجب أن نكون على دراية ومعرفة بالطرق والمناهج التي تمكننا من تعميم نتائج الدراسات التي تُجري على العينات لتشمل المجتمع الدراسي بأسره.

يتناول الفصل الثاني، دارس العينات، وإجراءات اختيارها. وبناءً على ما اكتسبناه من معرفة، خلال الفصل الأول والثاني، نتجه في الفصل الثالث إلى دراسة الإحصاء الاستنتاجي (2)، ونغطي بعض المفاهيم الأساسية المتصلة بالإحصاء الاستنباطي (3) ونبدأ في هذه المرحلة باستخدام المناهج الإحصائية، لتعميم النتائج، واختبار الفروض الإحصائية بأسلوب منطقي استقرائي.

نقوم في الفصل الرابع بالتوسع في استخدام تقنية الاستقراء، ونبدأ في اختبار العلاقات المتبادلة بين المتغيرات ذات الأهمية الخاصة لبحثنا، ويساعدنا استعراض المعايير التي تحكم العلاقات المتبادلة، في تلخيص المعلومات المتعلقة بشكل الارتباط بين المتغيرات المستخدمة في البحث وقوتها، ومن المهم أن نذكر، أن كل التقنيات التي نوقشت في الفصول السابقة تتصف بالفعالية والجدية، مما يستدعى لتطبيقها، استخدام معايير صارمة، وفروض جادة، فيما يتعلق بالبيانات ومناهج جمعها.

يُعد الموضوع النهائي المطروح في الفصل الخامس في غاية الأهمية، إذ يتعلق بمناهج المعايير للكميات غير المعلمية، ومن خلال دراستنا وتعرفنا على هذه المناهج، نصبح قادرين على إضافة شيىء من المرونة لقدرتنا البحثية.

نطرح من خلال هذا العمل، أمثلة مستندة على قضايا تهم المكتبيين وأخصائي المعلومات أو نضمن تفسيرات للنتائج التي توصلنا إليها، بغية توحيد إجراءات مشكلة البحث، ثم إضافة تمارين بسيطة وإجاباتها في نهاية كل فصل، حتى يتمكن القارىء من التطبيق الفوري، لما قرأه، وتعلمه من الدراسة النظرية. علاوة على ذلك، فقد تم وضع «قائمة شارحة للمصطلحات في نهاية البحث، قمنا فيها بشرح، وتوضيح

(2) Inferential Satistics

(3) Inductive Statistics.

وتعريف المصطلحات والمفاهيم الأساسية التي استخدمت في هذا العمل.

نأمـل أن تقـوم الأفكـار التي طرحنـاهـا في العمل، بإثارة الاهتمام بمجال البحث والإحصـاء، والتـزود برؤية جديدة لتفهم الأسئلة وصياغتها، والحصول على إجابات عليها مما يثير اهتمامات المتخصصين في مجال المكتبات والمعلومات.

تقديم المترجمين

يعاني تخصص المكتبات والمعلومات من نقص واضح في الانتاج الفكرى في مجال المناهج الاحصائية للمكتبين. وتطبيقاتها في المكتبات ومراكز المعلومات وشبكاتها. كما يفتقد التخصص الإطار النظري الرياضى المتطور الذي يساعد على تفسير نشاطات هذا التخصص وشرحها. وجرت العادة في معظم أقسام المكتبات والمعلومات في الجامعات العربية على أن توكل مهمة تدريس الأحصاء لطلاب المكتبات والمعلومات إلى أساتذة الاحصاء في كليات العلوم أو كليات التجارة والعلوم الادارية. وكان المنهج الدراسى في هذا المقرر يقتصر على بعض مبادىء الاحصاء الوصفى ودون الدخول في أي نوع من التطبيقات المستثمرة من تخصص المكتبات والمعلومات. هذا إن وجد مثل هذا المقرر أساساً.

وكما هو معروف فإن تدريس الإحصاء لم يعد مقتصراً على تخصص بعينه، بل أصبح خادماً لكل التخصصات وطابعاً مميزاً للمنهج العلمي الذي يأخذ به الباحثون في كل العلوم: بحتة وتطبيقية واجتماعية وإنسانية، حتى صار من الصعب على أي متخصص في واحد من هذه العلوم فهم مايُنشر في الدوريات المتخصصة إن لم يكن ملماً ـ على الأقل ـ بالمبادىء الأولية لعلم الاحصاء. وتخصص المكتبات والمعلومات في أشد الحاجة إلى الاحصاء، ليس فقط في مجال البحوث التي تقود إلى معارف جديدة، ولكن أيضا في مجال استخدام البيانات التي تفسر الظواهر المعلوماتية، فالمكتبى أو رجل المعلومات يُعد مستخدماً لشتى أنواع البيانات، ويستفيد من دراستها في تطوير ورفع كفاءته المهنية لتقديم خدمات معلومات أفضل للمؤسسة الأم التي تنتمى إليها المكتبة، أو للمجتمع الذي تقوم بخدمته، وذلك لأن البيانات الإحصائية أصبحت جزءاً مهماً من الإنتاج الفكرى في المجالات التي يتعامل معها المستفيد من المكتبات ومراكز المعلومات.

ونحب أن ننوه أن الترجمة لم تكن ترجمة حرفية، بل هي أقرب إلى ترجمة المفاهيم منها إلى ترجمة النصوص. ولقد راعينا صعوبة المصطلحات في مجال الإحصاء ـ خاصة على طالب المكتبات والمعلومات المبتدىء ـ فقمنا بشرح بعض المفاهيم شرحاً موجزاً خارج

نطاق الترجمة. وحرصنا أن نضع الرموز في المعادلات الرياضية كما اصطلح عليها بين الرياضيين والإحصائيين سواءً كانت بالحرف العربي أو الروماني.

يتكون الكتاب من خمسة فصول أو جزها المؤلفان في مقدمتهما، وزُوِّد الكتاب بملحق يتضمن جدول الأرقام العشوائية، وجدول المساحات تحت المنحنى الطبيعي، وجدول توزيع t، وجدول توزيع مربع كاى X، وجدول معامل الارتباط، وجدول القيم الحرجة لـ R في اختبار Runs. وقد زوِّد الكتاب أيضا بتمارين في أخر كل فصل لتساعد الطالب على فهم محتوى الفصل، وقد تم حل هذه التمارين في آخر الكتاب ليتأكد كل طالب من سلامة فهمه وصحة حلوله. وقد زودنا الترجمة بقائمة للمصطلحات الاحصائية التي وردت في الكتاب مع شرح موجز لها تضمنت حوالى السبعين مصطلحاً، وقائمة أخرى بالمختصرات، وقائمة ثالثة بالرموز مع شرح لكل منها. وأخيراً قمنا بعمل كشاف موضوعى للكتاب كى يساعد قارئه على الرجوع بسهول ويسر إلى موضوع محدد في متن الكتاب.

نأمل أن نكون بهذه الترجمة قد أضفنا إلى المكتبة العربية مرجعاً نشعر بالحاجة الماسة إليه ـ خاصة لطلاب المكتبات والمعلومات، ولزملائنا من المكتبيين العاملين في المكتبات ومراكز المعلومات، وبذلك تكون هذه الترجمة قد ساهمت في تأصيل مفهومٍ واضح لعلم الاحصاء والتحليل الإحصائي، ومدى فائدتهما لتخصص المكتبات والمعلّومات. وكما ذكر المؤلفان في مقدمتهما أنه قد روعى خلال استعراض هذه الموضوعات التعامل مع الحد الأدنى المطلوب للمعطيات الرياضية، إذ ليس من أهداف الكتاب ولامن أهداف ترجمته أن نؤهل القراء ليصبحوا أخصائيين متخصصين في الاحصاء، ولكن الهدف هو إحاطة طلاب المكتبات والمعلومات والمكتبين وأخصائي المعلومات بالتقنيات الاحصائية الأكثر شيوعا واستخداما مع تطبيقاتها في تخصص المكتبات والمعلومات.

نأمل أن تحقق هذه الترجمة الأهداف المقصودة منها، وأن يلتمس لنا إخواننا الإحصائيون، وإخواننا المكتبيون أي أخطاء، وياحبذا لو نبهونا لها لتصحيحها في الطبعات القادمة، والله وراء القصد.

المترجمان

سيد حسب الله   محمد جلال غندور

الفصل الأول

الاحصاء الوصفـــي   DESCRIPTIVE STATISTICS

الفصل الأول
الاحصاء الوصفي
DESCRIPTIVE STATISTICS

يمكن تصنيف الإحصاء إلى نوعين رئيسيين، الإحصاء الوصفي، والإحصاء الاستنباطي (الاستقرائي). ونستطيع عن طريق استخدام الإحصاء الوصفي، تلخيص كم كبير من البيانات التي تحتوي على «متغير» أو «عدة متغيرات» مما يساعدنا على تفهم تلك البيانات وإبرازها بطريقة أكثر وضوحاً. ونعني هنا بمصطلح «متغير» أي بيان، أو قيمة، أو عملية، أو مفردات، أو أي ظاهرة أخرى يُراد تحليلها، ومن الجلي أن أي ظاهرة قابلة للتحول أو التغيير يمكن أن تدعى «متغير» ومن المؤكد أيضاً، أن المتغيرات في أي نوع من أنواع التحليل الإحصائي، يجب أن تتميز بالفعالية والنشاط والقابلية للقياس.

فمثلاً، إذا كنا نهتم بمدى الاستفادة من مقتنيات مكتبة ما، فإن أحد المتغيرات القياسية التي يجب أن تُختار لتحليل هذه الاستفادة، يمكن أن يتمثل في «تداول مقتنيات المكتبة»، على اعتبار أن سجلات التداول متاحة، أو على الأقل يمكن عملها والحصول عليها. في هذه الحالة، لو كان لدينا السجلات الإحصائية الخاصة بالتداول اليومي للمقتنيات، والتي تغطي فترة عام كامل، فسوف نكون بحاجة للقيام ببعض العمليات، لنجعل هذه الإحصائيات مفهومة، وذات فائدة، وقابلة للتحليل، وفي هذا الصدد يمكننا ـ مثلاً ـ إيجاد متوسط التداول اليومي للمقتنيات لمدة عام، وتُعد إجراءات القيام بمثل هذه العمليات نوعاً من أنواع التلخيص الإحصائي وفي هذا المستوى من التحليل الإحصائي الوصفي، نحن نحصر اهتماماتنا واستنتاجاتنا في الحالة البحثية التي يتوافر لدينا بيانات كاملة عنها.

إما في حالة امتلاكنا لجزء فقط من البيانات الكاملة ـ التي تمثل متغيراً أو أكثر ـ لحالة البحث، فإننا نلجأ عندئذ إلى اتباع إجراءات البحث الإحصائي الاستنباطي، أو

الاستقرائي، وبتعبير آخر، إذا ما أردنا الخروج بمؤشرات عامة موثوق بها، حول عملية «تداول مقتنيات المكتبة» ولم نجد بين أيدينا إلا كم محدود من المعلومات التي تمثل مجرد عينـات من التـداول اليومي لمقتنيات المكتبة، فإن المعايير الواجب اتباعها ـ في هذه الحالة ـ تختلف اختلافاً كلياً عن تلك المُتبعة في الإحصاء الوصفي، وتتصف مجموعة المعـايير هذه بالمنـطقية والـدقة العلمية. وسيتم مناقشتها في الجزء المتعلق بالإحصاء الاستنباطي في هذا العمل.

مستويات القياس Levels of Measurement

إننا نستخدم مصطلح «قياس» بدون تحديد دقيق له، كما لو أنه مصطلح واضح لا يتسم بالغموض، وعادة لا يمثل مفهوم مصطلح «قياس» أي مشكلة، خاصة إذا كنا نتعـامـل مع مفـاهيم المقـاييس المعيارية، أو وحدات القياس المتعارف عليها، مثل «القدم»، «البوصة»، أو النظام المتري. ومع ذلك، فالعديد من المتغيرات ذات الأهمية القصوى لمجال المكتبات غير قابلة للقياس بهذه المعايير، والإجراءات الإحصائية ترتكز على المنطق الرياضي الذي يتضمن فروضاً مؤكدة حول مستويات القياس.

لن نستعـرض هنـا المبادىء الرياضية للإجراءات الإحصائية المستخدمة في هذا الكتاب، ولكن يجب علينا ـ على الأقل ـ أن نفهم متطلبات الاختبارات المحدودة، والإجـراءات الأخـرى التي سنتعامل معها من خلال هذا العمل، والتي تستند على الفـروض والمبـادىء الـرياضية. ويرينـا الجـدول رقم (1) الاختبـارات والقياسـات الإحصـائية المناسبة، التي يمكن تطبيقها على مستويات القياس الأربع: الإسمى، الترتيبي، الحدى، النسبي، والتي سنقوم باستخدامها في هذه الدراسة.

الإسمـــــي، Nominal

لا يُعـد مستـوى القياس «الإسمى» بالشيء الجـديد على المكتبيين وأخصـائي المعلومات. حقيقة، إن خطة التحليل الأساسية لأي من العلوم، تُعد خطة إسمية، بمعنى أنها خطة تصنيفية، وقد يبدو ترتيب الأشياء (الفصل مابين الأشياء) حسب النـوع عملية بديهية وواضحة، إلا أنها بالرغم من ذلك عظيمة الأهمية، وهي كأي معرفة بشرية، تبدو سهلة، بعد أن يوضع لها نظام ومنهجية ويتم تطبيقها.

يُعد نظام ديوي العشري، مثالاً واضحاً للمستوى الإسمي للقياس، فهو يُفرق بين الكتب أو بين أوعية المعلومات الأخرى بمقياس مدى ملاءمتها لرقم تصنيف معين. كما أنه بإمكاننا قياس جمهور المكتبة إسمياً. بتصنيفه لعدة فئات: فمثلاً يُصنف إلى

جدول (1): الاختبارات والقياسات الإحصائية لتحليل البيانات

القياس الإسمي Nominal

متغير واحد	متغيران
لحساب الـ التناسب Proportions النسب المئوية Percentages معدلات التغير Ratios نسبة (نسب) التغير Rates of Change المنوال Mode	اختبار فروض مربع كاي (Chi-Square) للمتغير المستقل 2(X) حساب الـ: معامل الصدفة Contingency Confficient (C) - قياس الارتباط معامل (T) Tschuprow'sT - قياس الارتباط - معامل (V) لكرامر Cramer's V - قياس الارتباط

القياسي الترتيبي Ordinal

متغير واحد	متغيران
المنوال Mode	اختبار Runs لوولد ولفيتز Wold-Wolfwitz Runs Test لحساب الـ: معامل الارتباط للترتيب النظامي (rs) : لسبيرمان Spearman rank - order correlation coefficient قياس الارتباط تو لكيندال (T) Kendall's tau measure of association (T)

القياس النسبي أو الحدي Ratio or Interval

متغير واحد	متغيران
اختبار الفروض على: المتوسط الحسابي لمجتمع الدراسة (اختبار t):(t test) حساب الـ: المتوسط (الوسط) الحسابي Mean المتوسط (الوسط) العددي Median المنوال Mode انحراف المتوسط الحسابي Mean deveation الانحراف المعياري Standard deviation معامل التغاير (التباين) Coefficient of Variability	اختبار الفروض على: معامل ارتباط مجتمع الدراسة: Population Correlation Coefficient (P) الفروقات بين المتوسطات الحسابية لمجتمعي دراسة (اختبار t): Difference between 2 population means (t test) حساب لـ معامل الارتباط الخطي (r) لبيرسون: Pearson product-moment correlation Coefficient (r)

مستفيدين وغير مستفيدين، وفئة المستفيدين تصنف بدورها إلى فئات فرعية، ذكور وأناث، أو موظفين وعمال. أما العاملين بالمكتبة، فعادة يتم تصنيفهم إلى ثلاث فئات، الكـوادر المؤهلة، الكـوادر شبه المؤهلة، والكوادر الكتابية، ومن المؤكد أن مثل هذه التصنيفات، تُعد من الخطوات التمهيدية الأولى لأي من مستويات القياس ولكنها في حد ذاتها ذات أهمية قصوى، فالبحث في الأصناف المتعددة، وإيجاد العوامل المشتركة بينها، يمكن أن يقودنا إلى فهم الكثير عن طبيعة هذه الأصناف، ونستطيع أن نتوصل عن طريق هذه التصنيفات، إلى إجابات لمثل الأسئلة التالية:

- أي الفئتين (المستفيدين، غير المستفيدين) تحتوي على نسبة أكبر من الذكور أو الأناث ؟
- أي الفئتين تحتوي على نسبة أكبر من خريجي الجامعات؟

من الطبيعي أن كل فئة من هذه الفئات تتضمن أرقاماً حقيقة، ومن خلال كل فئة يمكن معالجة هذه الأرقام بطرق مختلفة. وتبرز مشكلة مستوى القياس عندما يوضع افتراض خاطيء تتم على ضوئه معالجة الفئات على أنها منتظمة ومتسلسلة كمياً، بينما هي في حقيقتها غير مترابطة وغير كمية (أي كيفية) (Qualitative) ويتم معالجة هذه المشكلة، وتـوضيح اللبس فيهـا عن طريق منـاقشـة المستـويات الأخرى للقياس، والتقنيات الخاصة لتحليل البيانات الإسمية.

الترتيبـــي Ordinal

يُعيننـا القياس الـترتيبي على التفـرقـة بين العـلاقات فيما يخص الترتيب النظامي، فمكتبـة ما قد تكـون أعـظم أو أقـل شأنـاً من مكتبـة أو مكتبات أخرى، فيما يتعلق بمتغـيرات معينـة، فمقتنيات المكتبـة (أ) قد تكـون أكـبر من مقتنيات المكتبة (ب)، ومقتنيات تلك الأخيرة قد تكون أكبر من مقتنيات المكتبة (ج). وعن طريق استخدام أسلوب القياس الـترتيبي، نستـطيع أن نتوصل إلى نتائج أكثر فعالية مما لو استخدمنا أسلوب القياس الأسمي، فباستخدام هذا الأسلوب (القياس الترتيبي)، نكون قادرين - ليس فقط - على استنتاج أن مقتنيات مكتبة ما متميزة منهجياً، ولكننا - أيضاً - نستطيع مقارنتها بمقتنيات المكتبات الأخرى، على الأقل فيما يتعلق بترتيبها حسب أهمية كل منها.

وعـادة ما تقاس الأفضليات بمعايير الأسلوب الترتيبي، فمثلاً، إذا استطلعنا آراء المستفيدين من مكتبة ما حول أهم ست خدمات رئيسية تقدمها هذه المكتبة، ونطلب

منهم ترتيبهـا حسب أهميتها، سنجد أن المستفيدين يقومون بعمل تلك القائمة المرتبة حسب وجهة نظرهم في الأهمية النسبية لتلك الخدمات، إلا أننا نكون على قناعة بعدم استطاعتهم تحديد الأسباب أو المبررات التي دعتهم إلى اختيار هذا الترتيب بدقة. ومثال لذلـك، إذ طلب من المستفيدين ترتيب «الخـدمـات المـرجعية»، «مجموعات الكتب» و«المواد السمعية والبصرية» حسب أهمية كل منها، قد يلجأ المستفيد إلى وضع «مجموعات الكتب» على رأس القائمة (أي الأكثر أهمية)، ويعقب ذلك بالمواد السمعية والبصرية، ثم يضـع في النهاية الخدمات المرجعية. بهذه الطريقة نستطيع ـ فقط ـ أن نتعرف على المـوقـع النسبي لكـل من هذه الخـدمات مقارنة مع الأخريات، ولكن لا نعرف مدى الاختلاف في الترتيب بين الأول والثاني، أو بين الثاني والثالث. ونسوق مثالاً آخر: إذا طرحنا سؤالاً، ليستفسر من المستفيد عن أفضل عشر قصص لديه؟ فقد نجد أن قصة «الحرب والسلام»، تجيىء في المقدمة بصورة كاسحة وملحوظة، بينما نجد أن الفروقات بين ترتيب القصص التسـع الأخـرى غير ملحـوظـة، بمعنى أن نتيجة التحليل تشير بوضـوح أن نقـاط تفـوق القصـة الأولى (الحـرب والسـلام) تبـز بكثير التسع قصص الأخـرى، بينـما نتـائج التحليل لا تبرر بوضوح فروقات الترتيب بين القصص التسع الأخرى. يرجع السبب في ذلك أن جمهور المستفيدين يكون لديه أفضلية أولى متميزة يضعها في بداية القائمة، يعقبها بإثنين أو ثلاث من العناوين المفضلة لديه عن غيرها، أمـا بعـد ذلـك، فيقـوم بترتيب العناوين بطريقة تتسم بعدم الشعور الجدي بأهمية أو أفضلية كل منها عن الأخرى.

الحـــدي Interval

يمكننـا القياس الحـدي ـ ليس فقط ـ من ترتيب الملاحظات عن المتغيرات حسب أهميتهـا، بل أيضـاً من وضـع «وحدة قياس ثابتة»، تستخدم في تحديد التدرج ونقطة الصفر العشوائية. ونعني «بوحدة القياس الثابتة»، أن المسافات بين النقاط المتجاورة في التدرج تكون ثابتة (أي متساوية).

ولشرح ذلك في مثال: يقاس المتغير (IQ) ـ عادة ـ بمستوى القياس الحدي، ويعني هذا أن المسافة بين 90 -100، تساوي نفس قيمة المتغير (IQ)، للمسافة التي بين 110-120 (هذا يعني أن قيمة المتغير (IQ) = 10 وحدات).

وإذا كان المتغير (IQ) يساوي صفراً في تجارب قياس مستوى الذكاء، فإن هذا لا يعني أن ذكاء الشخص الخاضع لتجربة قياس الذكاء يساوي صفراً، طالما أن اختيار موقع

المتغـير (IQ) من التـدرج تم عشـوائياً (أي باستخدام نقطة الصفر العشوائية)، وذلك انـطلاقـاً من الحقيقية التي تنص على أن الشخص المتـوسط الـذكاء يمكن أن يحصل ـ بسهولة ـ على 500 درجة. يتضح من هذا، أن الأرقام قابلة للتفسير على ضوء الأداة (المتغير) (IQ)، الذي يستخدم في المقام الأول لقياس المفاهيم، وبعض المتغيرات التي لها أهمية خاصة للمكتبيين، تقاس باستخدام «المستوى القياسي الحدي».

النسبـــــة Ratio

عنـدمـا تكون المسافات بين النقاط في الترتيب المتدرج، محددة ومعروفة، وتتضمن الصفـر المطلق، فإن كثيراً منا يأخذ مسألة تدرج النسبة كمعيار للقياس، قضية مسلماً بها. وعـادة ماتكـون وحـدات القياس لمثل هذا الترتيب، ثابتة ومعيارية، وتسمح لنا بإجراء تحليلات دقيقة للغاية. مثال لذلك: إذا كان فرع مكتبة (أ) يحتوي على أرضية مساحتها 5000 قدماً مربعاً، وفرع مكتبة (ب) يحتوي على 4000 قدماً مربعاً، بينما فرع مكتبة (ج) يحتوي على 2000 قدماً مربعاً هذا يعني، أن الفرع (ب) أصغر من الفرع (أ) وأكبر من الفرع (ج) (مما يعني، أننا نستطيع ترتيب هذه الفروع بطريقة متدرجة بناء على مساحة كل منها). بالإضافة إلى أننا نستطيع ـ أيضاً ـ تحديد مدى الفروقات بين هذه الفروع بطريقة دقيقة سواءً بالأقدام المربعة أو بالنسبة المئوية التي تمثلها هذه الفروقات.

فالفرع (أ) أكبر من الفرع (ب) بألف (1000) قدم مربع، أي ما يساوي 25%
والفرع (أ) أكبر من الفرع (ج) بثلاث آلاف (3000) قدم مربع، أي ما يساوي 150%
والفرع (ب) أكبر من الفرع (ج) بألفي (2000) قدم مربع، أي مايساوي 100%

والقيمة صفر ـ في هذه الحالة ـ تعني، عدم وجود أي مساحة مطلقاً، وتشير إلى الغياب الكامل لأي مساحة تُقاس، ومن هنا نطلق مايسمى بنسبة «الصفر المطلق».

يصبو المرء دائماً، في الحالات التي يقوم بدراستها، أن تكون قياساته أقرب ماتكون إلى الدقة، إضافة إلى أنه يسعى إلى قبول واستخدام أسلوب «القياس النسبي» لنوعيات من القياسـات تحتـاج إلى حرص ودقة كبيرتين لتفسيرها وتحليلها، ولنأخذ مثالاً لذلك يتعلق بالمقـارنة بين ميزانيات المكتبات، لو افترضنا أن ميزانية المقتنيات لمكتبة جامعية تقدر بـ 600,000 دولاراً، ومكتبة أخرى ميزانيتها 400,000 دولاراً، وثالثة تقدر ميزانيتها بـ 200,000 دولاراً. يمكننـا في هذه الحـالة تقدير الفروقات بين الميزانيات الثلاثة، بعدد الدولارات الفعلية، وكذلك بالنسب المئوية بين تلك الميزانيات، ويتم ذلك بدون أي

صعوبة تذكر ، ويمكننا القول في هذه الحالة ، بأن المكتبة الأولى أغنى من المكتبة الثانية بنسبة 50% ، وأغنى من الثالثة بنسبة 300% أي ما يعادل ثلاثة أضعاف المكتبة الثالثة. ولكن ماذا يُفهم من التصريح بأن مكتبة ما ، تمتلك ثلاثة أضعفا ميزانية مكتبة أخرى ؟ فالميزانيات تماثـل في مفهـومهـا دخل الفرد. فهي وإن كانت تُقاس بالدولار ، إلا أنها تخضـع لمعايير ومقاييس تتعلق بقيمة الدولار، وقوته الشرائية ، وما يمكن أن يفعله من أجل هذه المكتبة. قد يتم شرح ذلك على النحو التالي :

إن مئتي ألف دولار تمكن المكتبـة من تدعيم المجمـوعـات المتـداولـة للمـطبوعات الأمـريكية. ومئتي ألف دولار أخرى تمكن المكتبة من تنمية مقتنياتها (من زاوية الكم والكيف). (مثال لذلك زيادة عدد عناوين الدوريات المشتركة فيها المكتبة)، وإضافة مئتي ألف دولار ثالثـة تمكن المكتبـة من زيادة مصادر المعلومات وتنويعها والدخول في مجالات جديدة (مثـلا: إضـافـة مقتنيات بلغات جديدة)، وفي الواقع العلمي تظهر الفـروقات بشكل واضح بين ميزانيات المكتبات، عندما توضع هذه الميزانيات موضع التنفيذ، فهي تسـاعـد المكتبات ذات الميزانيات الضخمة على الحصول على مقتنيات جيدة ومتميزة، وتقديم خدمات المعلومات بمستوى أفضل وأكثر جودة. ويكن تعميم هذه النتيجة ، إذا كنا بصدد مقارنة الدخل الذي يحصل عليه المكتبيون أو رواد المكتبة ، وعلاقة ذلك بالخدمات المكتبية ، فإذا اعتبرنا أن مبلغ (8000) دولار، الدخل الأساسي السنوي لأسرة من الطبقة المتوسطة مكونة من ثلاثة أفراد، ورأينا أن نضاعف دخل هذه الأسرة ـ وإن كان ذلك ليس بالأمر السهل ـ فسوف يغير ذلك من أسلوب معيشتها ، وقد يؤثـر على تعاملها واستخدامها للخدمات المكتبية. ومعرفة هذه الحقائق ، لايعفينا من إجـراء تحليلنـا باستخدام النسبة المئوية «أو أي من أشكال القياس النسبي»، إلا أنها تجعلنا نأخذ بعينه الاعتبار الآثار المترتبة على ذلك عند تحليل وتفسير البيانات اعتماداً على هذا الأسلوب.

يشكل لنا العرض السابق قاعدة تمهيدية عن مناهج استخدام الأرقام في التعرف على قيم وخصائص المتغيرات ، ذات الاهتمام الخاص بمجال المكتبات والمعلومات ، وسنقوم فيما تبقي لنـا من هذا الفصـل ، بتقييم بعض الإحصاءات المنتقاة ، المتعلقة بتوصيف البيانات الإحصائية وتفسيرها.

شرحنـا سابقـاً أن الإحصـاء الـوصفي ، يمثل تلخيصاً للمعايير (المقاييس)، التي تسـاعـدنـا في بلورة الكم الكبيـر من المعلومـات المتعلقـة بالحالات والموضوعات التي

ندرسهـا، ونحيلهـا ـ أي المعلومـات ـ إلى مجمـوعة محدودة من الأرقام التفسيرية ذات المغـزي لدراستنـا، وسوف نبدأ بدراسة بعض المقاييس المألوفة التي وردت مرتبطة مع بعضهـا البعض، وهي بالتحديد، النسبة المئوية "Percentages" النسب "Rates"، المعدلات Ratios وسنقوم بعد ذلك بمناقشة بعض المقاييس الجديدة، التي تفيد بتعريف معدلات متوسط المجموعة، وتبرز المعلومات التي توضح كيفية تشابه الحالات الفردية التي ندرسها، كما أننا سوف نشير إلى علاقة الحالات الفردية بالمجموعة ككل، من خلال التعرف على مواقعها وترتيبها داخل المجموعة.

تلخيص الإحصاءات Summarizing Statistics

تُعد بعض السبل المفيدة لتلخيص البيانات مألوفة لدرجة أنها ـ غالباً ـ لا تلاحظ، مثال لذلك التناسب Proportions، النسبة المئوية، النسب، والمعدلات، وهي ذات فائدة كبيرة لأنها تتيح لنا مقارنة المجموعات ذات الأحجام المختلفة، عن طريق توحيد وحدات القياس لهذه الأحجام.

التناســـب Proportions

عنـدمـا أثير تساؤل حول التشدد النسبي فيما يخص اقتناء الكتب المرجعية المكتوبة بلغات غير اللغة الإنجليزية بإحدى المكتبات، أجريت دراسات على مجموعات المراجع في مجالات العلوم الإنسانية والعلوم الاجتماعية بالمكتبة الجامعية، وقد اتضح أن مجموعة العلوم الإنسانية تحتوي على (2360) عنوانا، منها (880) عنواناً بلغات غير الإنجليزية، أما مجموعة العلوم الاجتماعية فوجد أنها تتألف من (1890) عنواناً منها (440) عنواناً بلغات أخـرى غير الإنجليزية. ويتضح من ذلك، أن التناسب مابين أعداد عناوين الكتب باللغـات الأجنبية مقـارنة بالمجموع الكلي لعناوين الكتب يساوي 2360/880 (بالنسبة للعلوم الإنسانية)، و 1890/440 (بالنسبة للعلوم الاجتماعية)، ونرى هنا أن مجرد تقسيم بسيط يُعبر عنه بعلامة خط فاصلة (/)، يُسهل مهمتنا للمقارنة بين المعلومات المتوافرة لدينـا عن المجمـوعتـين. والتنـاسب الـذي تم استنتـاجه من المعلومات المُعالجة بهذا الأسلوب يساوي 37, و 23, على التوالي [على فرض أن القياس لكل العناوين يساوي 1,0 في كل مجموعة].

في مثال آخر، نجد أن من بين كل 1000 مستفيد (لا متخرج) في مكتبة جامعة ما، يوجـد (250) مستفيداً مسجلين في برامـج دراســات مستقلة، و (750) مستفيداً غير مسجلين في أي برامج، إذاً نسبة المستفيدين المسجلين في الدراساته المستقلة يساوي

1000/750 ، أي ما يقدر بـ 75, .

ومثال أخير، وُجد أن من بين (150) كتاب قصص تم تداولها في المكتبة، هناك (100) قصـة للبـالغـين، و (50) قصـة للصغار، ونستنتج من ذلك أن نسبة تداول القصص المصنفة للكبار بلغت 150/100 ، أي ما يساوي 67, ونسبة القصص المتداولة والمصنفة للصغار تبلغ 150/50 ، أي مايساوي 33,

النسبـــة المئويـــــة Percentage

يتم ـ ببساطة ـ حساب النسبة المئوية، بمضاعفة التناسب مئة مرة (أي أجراء عملية ضرب التنـاسب في 100) ، وبـالنسبة للمثال الأول الذي أوردناه قبل قليل، نستطيع القـول بأن 37% من عنـاوين المـراجـع في مجال العلوم الإنسـانية مكتـوبة بلغات غير الإنجليزية، وأن 23% من المراجع في مجال العلوم الاجتماعية مكتوبة بلغات غير اللغة الإنجليزية. وفي المثال الثاني 25% من المستفيدين بخدمات المكتبة ملتحقين بدراسات مستقلة، بينما 75% غير ملتحقين بأي دراسات. أما المثال الثالث، فـ 67% من القصص المتداولة خاصة بالكبار، بينما 33% من القصص المتداولة خاصة بالصغار.

حقيقـة، أن معـدل استخـدام النسب المئـوية، أكـبر بكثـير من معـدل استخدام التناسب، ولكن كلاهما يُعتبر مألوفاً ومستخدماً بما فيه الكفاية، ليجعلنا في حاجة دائمة لتذكر أوجه التشابه والاختلاف بينهما.

في جدول رقم (2) نجد أن مجموعة المكتبة الخاصة، تحتوي على عناوين قصصية تزيد بنسبة 33% عن مثيلتها في مجموعة مكتبة المدينة، ولكن مكتبة المدينة تزيد بـ 1472 عنواناً قصصياً عن المكتبة الخاصة، وذلك إذا أخذنا بالمجموع الكلي لعناوين القصص في كل منها، ولذلك فإن الاقرار بالنسبة المئوية أو التناسب بدون الأخذ بالحجم الكلي للعينة قد يقود إلى تضليل القارىء.

جدول رقم (2) بيانات عن مجموعات المكتبات

مجموعة مكتبة المدينة			مجموعة المكتبة الخاصة		
	عدد	نسبة مئوية		عدد	نسبة مئوية
قصصية	1500	60	قصصية	28	93
غير قصصية	1000	40	غير قصصية	2	7
المجموع	3500	100	المجموع	30	100

ومن الضروري أن تتضمن الأمثلة التي نستشهد بها عدد حالات البحث، أو عدد الملاحظات، أو عدد العناوين في مجموعة المراجع، أو عدد المستفيدين، أو العدد الكلي للمراجع المتداولة، إذ أن هذا يجعلنا قادرين على معرفة القيمة الفعلية والحقيقية لأرقام النسبة المئوية والتناسب الواردة في الدراسة.

النسب والمعدلات Ratios and Rates

يمكن التعبير عن النسب بطرق مختلفة، ففي يوم عمل من أيام المكتبة، قد نكون أعرنا 150 كتاباً قصصياً، و300 كتاب غير قصصي، وبذلك يمكن التعبير عن نسبة إعارة الكتب القصصية إلى الكتب غير القصصية كالآتي:

300/150 أو 300:150 أو 150 إلى 300، المصطلح الأساسي في هذه العملية هو كلمة «إلى: To» أي كان الرقم الذي يرد قبل كلمة «إلى» فهو يقسم على الرقم الذي يليه. وعادة مانكون شغوفين بتبسيط الأشياء واختصارها، وعليه، فإن النتائج يمكن أن تكون على النحو التالي:ـ

30/15 أو 10/5 أو 30:15 أو 10:5.

وهدفنا ـ في الغالب ـ هو المقارنة بين النسب، والاستمرار في عمليات الاختصار، ويقودنا هذا إلى المرحلة التالية التي يطلق عليها «الوحدة Unity»، فإذا أخذنا المثال السابق، سنجد أن النسبة بين الكتب القصصية إلى الكتب غير القصصية يساوي 2:1 أو 1 إلى 2 (ظهور رقم أقل من واحد صحيح في النسب، يعد أمراً غير طبيعي)، والعلاقة نفسها ما بين الكتب غير القصصية إلى الكتب القصصية يمكن أن يُعبر عنها بـ 1:2، ومن الواضح أنه يمكن ـ بسهولة ـ الخلط ـ خطأ ـ مابين علاقات التناسب وبين النسب أو العكس. ولتوضيح ذلك نأخذ المثال السابق: حيث العدد الكلي للعناوين المتداولة يساوي 450 عنواناً، وتناسب الكتب القصصية يساوي 450/150، أي ما يساوي الثلث. تناسب الكتب غير القصصية يساوي 450/300 أي يساوي الثلثين. إذن نسبة الكتب القصصية إلى الكتب غير القصصية يساوي 2:1 كما هو موضح عاليه.

في مثال آخر: تم إعارة 100 كتاب إلى أشخاص بالغين، و50 كتاب إلى أطفال، إذن نسبة إعارة كتب البالغين إلى كتب الأطفال تبلغ 50:100، أي 1:2 إن مفهوم المعدلات مرتبط إلى حد بعيد بمفهوم النسب، ويستخدم غالباً لتجنب التعامل مع الأرقام العشرية الصغيرة (العلامات العشرية)، وبالفعل تظهر أهمية الأخذ بالمعدلات، عندما نتعامل مع أرقام كبيرة تتكون من المئات والآلاف. يمكن مثلاً: التعبير عن الكتب التي

فقدت نتيجة لإجراءات الإعارة، بالقول أن الكتب المفقودة تمثل نسبة من الألف، أو نسبـة من العشرة آلاف. وبـما أن الكثـير من المقاييس التي نستخدمها تتعلق بظروف متغيرة، فإن استخدام «معدل التغير» يكون ضرورياً وهاماً لنا، فمعدل التغير هو شكل من أشكـال النسب، ويمكن حسـابـه بإيجـاد الفرق مابين قيمة المتغير عند بداية فترة معينة، وقيمته عند نهاية هذه الفترة، ثم نقوم بقسمة قيمة الفرق على قيمة المتغير في بداية الفترة، مثال لذلك:

إذاكانت مقتنيات المكتبة عام 1965 تبلغ (50000) مجلداً، وبلغ عددها في عام 1975 (150,000) مجلداً، فإن معدل التغير، ينتج عن طريق قسمة الفرق مابين مقتنيات المكتبة عام 1965 ومقتنياتها عام 1975 أي (100,000) مجلداً على مقتنيات المكتبة في بداية الفترة (1965) أي (50,000)، وتظهر المعادلة كما يلي:

$$\frac{\text{عدد المجلدات عام 1975 ـ عدد المجلدات عام 1965}}{\text{عدد المجلدات عام 1965}} = \frac{\text{الفرق}}{\text{عدد المجلدات عام 1965}} = \frac{\text{رقم}}{\text{رقم}} = \text{رقم} = \text{الرقم}\ \%$$

أي

$$\frac{150000-50000}{50000} = \frac{100000}{50000} = 2 = 200\%$$

إذاً مجموعة هذه المكتبة تضاعفت بما يساوي 200%، للفترة مابين 1965 و 1975 (مع ملاحظة أن «معدل التغير» يمكن أن يظهر بقيمة سالبة، إذا اتضح أن حجم المقتنيات تناقص بدلا من أن يتضاعف).

نجد أن البعض يفضل أن يصف الحالة التي عليها مقتنيات المكتبة ـ في المثال السابق ـ بطريقـة مختلفة، وذلك بقولهم، أن مقتنيات المكتبة عام 1975 تساوي ثلاثة أضعاف ماكانت عليه 1965، ولكن هذا القول لا يمكن ـ بأي حال من الأحوال ـ أن يحول إلى نسبة مئوية، حيث أننا لا نستطيع القول بأن المقتنيات زادت بنسبة 300%. لذلك يجب أن نكون حريصين في الأخذ بما يظهر في بعض التقارير، حيث أن عملية حسابية كهذه، يمكن أن تقودنا إلى أرقام غير صحيحة، وبالتالي إلى استنتاج خاطىء، وعلى أي حال، أن مصطلح «معدل التغيير» يعبر تماماً عن مفهومة، أي أنه مجرد إجراء يقرر «التغير في المعدل».

ونورد هنا، مثالاً أخيراً لتأكيد هذا المفهوم: دعنا نفترض أنه تمت ترقيتك إلى منصب

مساعد المدير، بمرتب قدره 15000 دولاراً سنوياً، وكان مرتبك السابق 12000 دولاراً سنوياً، إذن معدل التغيير في مرتبك، يمكن حسابها كالآتي:

$$\frac{15000-12000}{12000} = \frac{3000}{12000} = ,25$$

$$= ,25\ \%$$ زيادة

معايير التلخيص Summarizing Measures

كنا نتعامل - حتى الآن - مع مجموعات من البيانات البسيطة نسبياً، والتي تحتاج إلى كثير من الجهد لمعالجتها وإعادة تنظيمها، ولنفهم إثنين من أهم معايير التلخيص يجب علينا استيعاب بعض المفاهيم المتعلقة بهذا المجال، وخاصة ما يتعلق بمفهوم «توزيع التكرار Frequency Distribution » والتكرار، يعني عدد المرات التي تظهر فيها قيمة معينة، فمثلاً، لو تحصل 5 طلاب في امتحان ما على 8,5 درجة لكل منهم، نستطيع القول، بأن تكرار الدرجة 8.5 يساوي 5. وعلى نفس المنوال، لو كان متوسط سن مديري فروع 322 مكتبة يتراوح ما بين 20 و 29 عاماً نستطيع القول، بأن تكرار المجموعة العمرية لمديري الفروع يساوي 322. والجدول الذي يرتب في شكل قائمة، بقيم المتغيرات، ويحدد التكرار لقيمة كل متغير للحالات موضع الدراسة، يسمى جدول «توزيع التكرار».

مثال لذلك، الجدول رقم (3) ، الذي يحتوي على بيانات تتعلق بعدد مرات التوقف التي تقوم بها، مكتبات متنقلة خلال رحلاتها في مناطق مختلفة. وفي هذا الصدد، يُعد «عدد مرات التوقف» لكل مكتبة متنقلة هو المتغير المتعلق بهذه المكتبة، وقد تم عرض هذه المعلومات في جدول (4) ، في شكل «توزيع تكرار» يوضح لنا عدد المكتبات التي توقفت 6 مرات، وتلك التي توقفت 9 مرات. . . الخ.

دعنا نفترض، بأننا التحقنا بمجموعة مكتبات تقوم بتقديم خدمات المكتبة المتنقلة في مقاطعات [1] متجاورة، وكان عدد المكتبات المتنقلة (وكذلك المقاطعات) يساوي 16 مكتبة (مقاطعة) وبانضمامنا لهذه المجموعة سيرتفع العدد إلى 17 مكتبة متنقلة - وبالرغم من ذلك سيظل المتغير الهام بالنسبة لنا هو «عدد مرات التوقف».

---

(1) مقاطعات: شكل من أشكال التقسيم الإداري للبلاد، يقابله في البلاد العربية مسميات مثل: المحافظات، المديريات، الإمارات (المترجم).

جدول (3): جدولة توقف المكتبات المتنقلة

الحـالة (المكتبة المتنقلة)	قيمة المتغير (عدد مـرات التوقف)
المكتبة أ	6
المكتبة ب	9
المكتبة ج	10
المكتبة د	14
المكتبة هـ	16
المكتبة و	17

جدول (4): توزيع التكرار، للتوقفات المجدولة

التكرار (عدد المكتبات المتنقلة)	قيمة المتغير (عدد مـرات التوقف)
1	6
1	9
1	10
1	14
1	16
1	17

الجدول رقم (5) يسجل أمام كل حالة بحثية (مكتبة متنقلة)، قيمة المتغير الخاص بها (عدد مرات التوقف) جدول (5) التوقفات المجدولة لـ 17 مقاطعة.

ويلاحظ أنـه، كلما زادت أعـداد الحـالات البحثية.في الـدراسـة، كلما زاد تعقيد الجداول التي تُعبر عنها، لذا، نحن في حاجة إلى طريقة أكثر إيجازاً لعرض المعلومات، تمكننا من جدولة المعلومات عن «نقاط التوقف» وبطريقة أكثر بساطة.

إذا كنـا نود ـ حقيقـة ـ عـرض المعلومات عـن نمط «نقاط التوقف» للسبع عشـرة مكتبة متنقلة بطريقة سهلـة، يمكننا استخدام نموذج « توزيع التكرار »، كمـا هـو موضح فـي جـدول (6).

جدول (5): التوقفات المجدولة لـ 17 مقاطعة

الحالة (مكتبة متنقلة)	قيمة المتغير (عدد مرات التوقف)	الحالة (مكتبة متنقلة)	قيمة المتغير (عدد مرات التوقف)
المكتبة أ	6	ي	10
المكتبة ب	9	ك	9
المكتبة ج	10	ل	14
المكتبة د	14	م	14
المكتبة هـ	16	ن	16
المكتبة و	17	ظ	9
المكتبة ز	14	ق	17
المكتبة ح	16	ر	16
المكتبة ط	14		

والملاحظ، أن المعلومات المعروضة فـي جـدول (6) ، هـي نفسها المعـروضة فـي جدول 5 ، إلا أن «التوزيع» فـي جـدول (6) معروض بطريقة أكثر بساطة وإيجـازاً، فمجالات البحـث (المكتبات المتنقلة) ـ هنا ـ معروضة فـي شكل مجموعات تضم كل منها الحالات المتماثلة في قيمة المتغير المعني بالبحث (عدد مرات التوقف).

ويُعتبر المضلع التكراري (Frequency Polygon) ، منهـج آخر ، مـن مناهج العرض المختصر لمعلومات كبيرة الحجم، تتعلق بمجموعة واحـدة مـن الحالات، يتم قياسها على ضوء متغير واحد. حيث يمكن عرض المعلومات المتعلقة «بتوزيع التكرار»، عن

جدول (6): «توزيع التكرار» للتوقفات الخاصة لـ (17) مقاطعة

المتغير (عدد مرات التوقف)	التكرار (عدد المكتبات المتنقلة)
6	1
9	3
10	2
14	5
16	4
17	2

طريق الـرسم البياني، وذلـك باستخدام «المضلع التكراري»، كمـا هـو موضح فـي شكل (1).

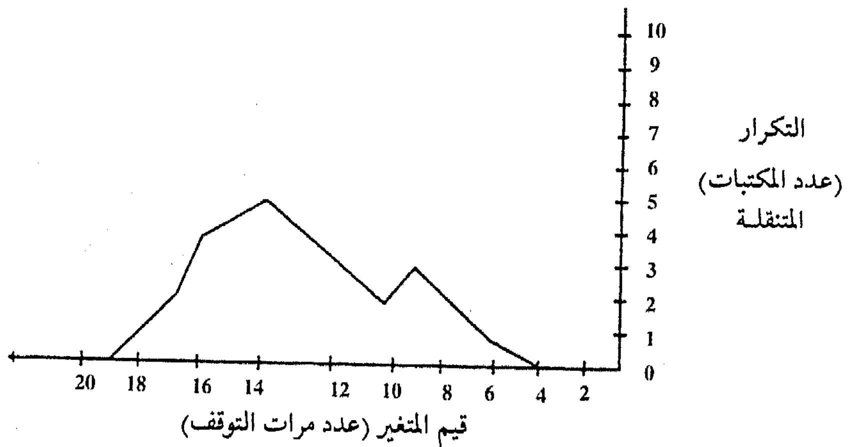
ولتـطبيق هـذا الأسلـوب ، نبـدأ برسم إحـداثيين ، الأفقي منهـا يمثـل القيـم المتعـددة للمتغير موضع البحث، أما الرأسي فيشير إلى تكرار الحالات لكـل قيمة مـن القيم المذكورة.

ونبدأ الرسم، بأن تحدد واحـدة مـن القيم للمتغير (وذلك على الإحداثي الأفقي)، ثم نبدأ بالصعود فـي خـط مستقيم إلى أعلـى - مواز للإحداثي الرأسي - ثم نضع نقطة أمام الرقم الذي يمثل تكرار الحالات لتلك القيمة.

(مثـال لذلـك: للعيـنـة التـي تساوي 6 توقفات) : نبدأ التحرك عـلى الإحداثي الأفقي حتى نصـل إلـى الرقـم 6. ثم نبدأ في التحرك إلى أعلى فـي خـط مواز للإحداثي الرأسي، ونتوقف أمام الرقم 1 ، ونضـع نقطـة . هـذا يعني أن عـدد المكتبات المتنقلـة التـي توقفت 6 مرات «يبلغ عددها» «مكتبة واحـدة فقـط».

وبهـذه الكيفية، نبـدأ فـي نقل الأرقام المسجلة فـي أعمدة «توزيع التكرار»، مـن الجدول إلـى الرسم البياني ، وبعـد أن يتم نقـل جميع الأرقام الدالة علـى التكرار إلـى «المضلع التكراري» بالرسم البياني، على شكل نقاط، نبدأ في رسم خط يصل ما بين هذه النقاط.

وبمجـرد نظرة خاطفـة إلى الـرسم الـذي يمثل «المضلع التكراري»، نستطيع أن نعـرف، بأن أكثـر القيم تكراراً في الظهور بلغت 14 مرة (5 مكتبات توقفت 14 مرة)،

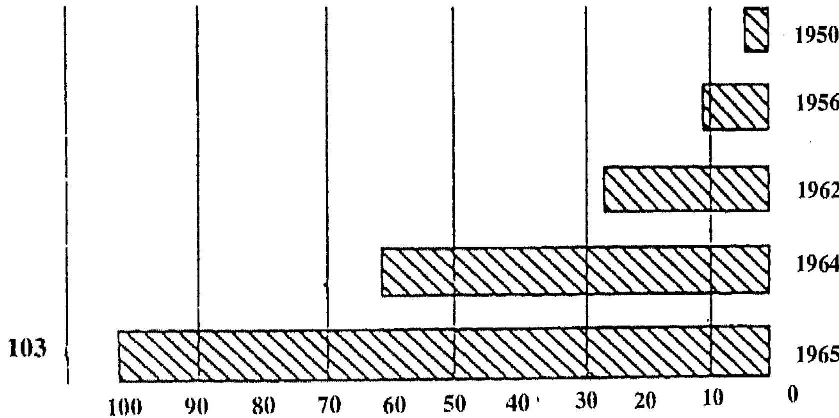


شكل (1): عرض مرات التوقف لـ 17 مقاطعة بالمضلع التكراري.

وبنفس الكيفية نستطيع التوصل إلى معرفة، أن أقل عدد مرات توقف بلغت 6 مرات، وأن أكثر عدد مرات توقف بلغت 17 مرة. يقوم «المضلع التكراري» بتحويل بيانات «توزيع التكرار» إلى شكل بياني، وهذا الأسلوب يسمح لنا برؤية البيانات بصورة شاملة، وبالتالي يمكننا فهم الموقف الكلي بطريقة أفضل.

شكل آخر من أشكال عرض المعلومات عن طريق الرسم البياني، «المدرج التكراري Histogram» ويعرف أيضاً «بالشكل البياني العمودي»، وفي هذا الشكل نعرض بيانات تكرار المتغير على الإحداثي الرأسي ويمكن ـ أيضاً ـ تسجيل بيانات التكرار على المستطيل الذي يمثل الحدث. (يمثل الشكل رقم (2) تطبيقاً لهذه التقنية).

إن بيانات «توزيع التكرار»، غالباً ما تكون كبيرة الحجم، وسواء تم عرضها ـ أو لم يتم ـ عن طريق تقنيات الرسم البياني، فإن حجمها هذا لا يسمح بأن يستوعبها القارىء من مجرد نظرة أو لمحة سريعة، وهناك تقنيات أخرى لعرض معلومات «التوزيع بطريقة أكثر اختصاراً»: الأولى تُدعى «مقاييس النزعة المركزية» "Measures of Central Tendency" والثانية «مقاييس التشتت "Measures of Dispersion" وبتعبير أكثر بساطة يعبر عن هذه المفاهيم بما يسمى المتوسط، أو الوسط، أو المُعدل، والتغير، أو الانحراف. وهذان الأسلوبان، ذوا فائدة جمة، في تلخيص كميات كبيرة من المعلومات، ووضعها بصورة مختصرة.



شكل (2): المدرج التكراري: ميزانيات (أوجه الصرف) للمكتبات العامة
(مقدرة بملايين الدولارات)

المصدر: Libraries at large, edited by Douglas M. Knight and E.A. Nourse (New York; Bowker 1969) p. 181.

النزعـة المركزية Central Tendency

سنعـرض في هذه الجـزئية إلى ثلاثة أساليب أو مقاييس، وهي : المتوسط (الوسط) الحسـابي Mean ، المتوسط (الوسـط) العـددي Median ، والمنوال Mode ، وثلاثتها معروفة باسم «المعدل المتوسط»، بالرغم من أن كل منها يتبع أسلوباً خاصاً للوصول إلى «المتوسط» .

يُعد «المتوسط الحسابي» هو الأسلوب الأكثر التصاقاً في مفهومه «بالمعدل المتوسط»، ويمكن حساب «المتوسط الحسابي» بإحصاء القيم الملاحظة للمتغير»، ثم تقسيم مجموع هـذه القيم على عـدد (الحـالات البحثية) . والمعـادلـة المستخـدمـة لحساب المتوسط الحسابي كالآتي :

$$\bar{X} = \frac{\Sigma x}{n}$$

حيـث :

Σ = قيمة الـ

X = قيمة كل حالة أو قيمة كل ملاحظة (أو مشاهدة)

n = المجموع الكلي لعدد الحالات، المفردات أو الملاحظات (للعينة)

$\bar{X}$ = المتوسط الحسابي للعينة

$$\bar{X} = \frac{10+8+6+4+2}{5} = \frac{30}{5} = 6$$

ويشير المثال السابق إلى وجود (5) حالات، قيمها على التوالي 10, 8, 6,4,2 وللتوصل إلى متوسط قيم هذه الحالات، يجب جمع قيمتها معا (2 + 4 + 6 + 8 + 10)، ثم قسمتها على عدد الحالات (5) ، والنتيجة التي نحصل عليها «المتوسط الحسابي»، ويساوي في هذا المثال (6) ، ويمثله في المعادلة الرمز $\bar{X}$. ولمزيد من الإيضاح، دعنا نحاول حساب «المتوسط الحسابي» لمقتنيات 7 مكتبات عامة :

الحـــالة (مكتبة)	قيمة المتغير (كتب)
أ	6200
ب	4300
ج	3800
د	7400
هـ	5100
و	6500
ز	5900

$$\bar{X} = \frac{5900 + 6500 + 5100 + 7400 + 3800 + 4300 + 6200}{7} = 5600 \text{ كتاب}$$

∴ المتوسط الحسابي لمقتنيات المكتبات السبعة = 5600 كتاب

أما «المتوسط العددي»، فهو تلك القيمة التي تتوسط «توزيع التكرار»، حيث كل القيم تكـون مرتبـة من الأعلى قيمة إلى الأقل قيمة (أو العكس)، ويكون عدد القيم (الحالات) أعلى «المتوسط العددي» مساو لعددها أسفله، ويرمز «للمتوسط العددي» بالرمز $\dot{X}$.

في المثال السابق للسبع مكتبات العامة، يتم استخراج «المتوسط العددي» عن طريق ترتيب مقتنيات المكتبات، بداية بالمقتنيات الأكثر عدداً، ونهاية بالمقتنيات الأقل عدداً (ويمكن ترتيبهـا عكسياً)، ثم اختيار الـرقم المقـابل «للحالة الأوسطية»، ويكون هو «المتوسط العددي»:

7400

6500

6200

5900 المتوسط العددي = $\dot{X}$

5100

4300

3800

إذا المتوسط العددي، للمكتبات السبعة العامة يساوي 5900 ، حيث:

3 مكتبات لديها عدد كتب أكثر من المتوسط العددي .

و 3 مكتبات لديها عدد أقل من المتوسط العددي .

مثال آخر ، إذا أخذنا الأعداد 2, 4, 6, 8, 20 يكون «المتوسط العددي» (6) ، ولكن ماهـو «المتـوسط الحسابي» لهذه الأرقام؟ . يبلغ المتوسط الحسابي لهذه الأرقام (8). يبدو واضحـاً أن «المتـوسط العدد» أقل تأثراً من «المتوسط الحسابي» بالحد الأقصى والأدنى للقيم ، ويعـطي «المتوسط الحسابي» بطبيعته ، فرصاً متساوية لجميع القيم ، حيث يتم عرضهـا بنـاء على تناسبها المباشر للوزن الحقيقي ، ولذا يُعد «المتوسط الحسابي» ، من وجهـة النظر الرياضية البحتة ، من أكثر الوسائل أمانة ومصداقية وتمثلاً لمفهوم «النزعة المـركزية» لأنه يأخذ في الاعتبار كل القيم المُمثلة في الحالة الدراسية . ومع ذلك ، فقد يحدث أحيانـاً بعض الخلط وعدم الفهم ، عندما يُستخدم «المتوسط الحسابي» كوسيلة لتلخيص البيانات ، كما حدث ورأينا في المثال الأخير.

تُعـد الأجور والدخل من جملة التغيرات التي تنحو في بعض الأحيان إلى الظهور، بطريقة محرفة ، أي بأكثر من قيمتها الحقيقية ، ويظهر ذلك في المثال التالي :

مجموعة أ			مجموعة ب		
حالة	دخل		حالة	دخل	
1	2000	دولار	6	2000	
2	4000	"	7	4000	
3	6000	" [المتوسط الحسابي]	8	6000	[المتوسط العددي]
4	8000	"	9	8000	
5	10000	"	10	22000	

في المجموعة (أ) كلا من المتوسط الحسابي والمتوسط العددي يساوي 6000 دولاراً

حيث المتوسط الحسابي $\frac{10000+8000+6000+4000+2000}{5} = \frac{30000}{5} = 6000$ دولاراً ،

والمتـوسط العددي في هذه الحالة رقم (3) يساوي 6000 دولاراً ، في حين أن المتوسط الحسابي يساوى 8400 دولاراً (يمكن حسابها بنفس الطريقة السابقة) .

إذاً بالنسبـة للمجمـوعة (ب) : بما أن 4 أشخاص من خمسة دخلهم أقل من «المتوسط الحسـابي» لمجمـوعتهم (حيث أن المتوسط الحسابي يساوي (8400) دولاراً ، وهناك (4) أشخـاص دخولهم 2000, 4000, 6000, 8000 ، اذا المتوسط العددى (6000) دولارا وهو

الوسيلة الأكثر تمثيلاً للدخل لهذه المجموعة، حيث يمثل قيمة أوسطية لقيم الدخول لهذه المجموعة. وذلك لأن توزيعات المرتبات، عادة، ماتكون كما هو عليه الحال في المجموعة (ب) ولذلك يعتبر أسلوب «المتوسط العددي»، صحيح من وجهة النظر الرياضية. ولذا يُفضل تطبيق وإجراء الأسلوبين معاً، وإظهار نتائجهما جنباً إلى جنب، خاصة إذا كانت هناك فروقات أساسية بين النتائج التي نحصل عليها من كل منهما، بدلاً من الحيرة والقلق في التعرف على أفضلهما لتلخيص المعلومات التي بين أيدينا، وبهذا سوف نزود القاريء بمعلومات كافية تساعده في التعرف على مغزى الأرقام المتعلقة «بتوزيع المرتبات».

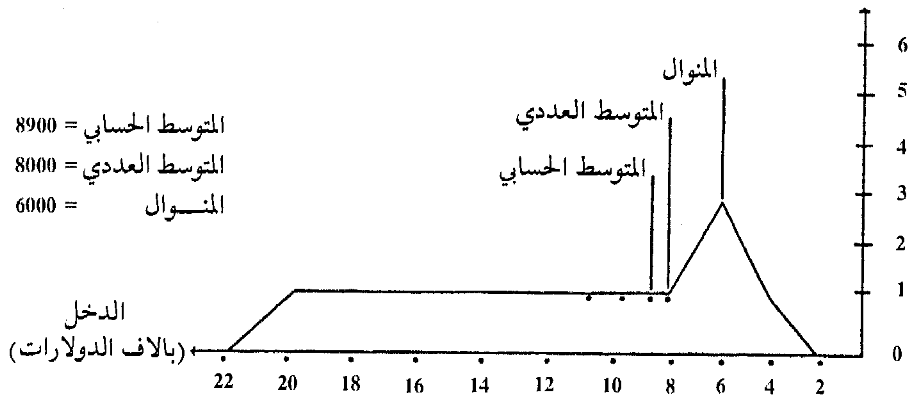
وعندما تكون عدد الحالات زوجية، نحتاج إلى إجراء عملية حسابية بسيطة، للتعرف على «المتوسط العددي» مثال لذلك: إذا أخذنا الترتيب العددي 12, 10, 8, 6, 4, 2 ، لانستطيع تحديد قيمة «المتوسط العددي» (أي عدداً تكون عدد القيم قبله يساوي عدد القيم بعده) حيث أن الأعداد زوجية ولحساب «المتوسط العددى» نقوم «بعملية جمع مابين «القيمتين الأوسطيتين» (6 + 8) ، ونقسم حاصل الجمع العددي على 2 ، لنصل إلى «المتوسط العددي» الذي يساوي في هذه الحالة 7.

أما المنوال (Mode) ، فتعني قيمة المتغير التي تظهر متكررة أكثر من غيرها (أي عدد مرات تكرار ظهورها أكثر من غيرها من القيم). في كل الأمثلة التي أوردناها سابقاً، لم يكن هناك ما يمثل مفهوم المنوال، فكل القيم ظهرت مرة واحدة فقط.

ولكن قد يحدث ان يكون لدينا (9) عاملين في مكتبة ما، دخولهم موزعة كالنحو التالي:

4100, 6000, 6000, 6000, 8000, 9000, 10000, 11000, 20000 دورلاً. المنوال ـ في هذه الحالة ـ يكون 6000 دولاراً ويرمز إليه بالرمز X.

ولكن، ماهي قيمة «المتوسط الحسابي» و «المتوسط العددي» في هذا التوزيع؟ـ الشكل رقم (3) يوضح مواقع المعدلات الثلاثة «المتوسط الحسابي»، «المتوسط العددى» «المنوال»، على ضوء علاقاتهم بعضهم ببعض.



شكل (3) عرض مرتبات (9) مكتبيين بالمضلع التكراري نموذج توزيع موحد

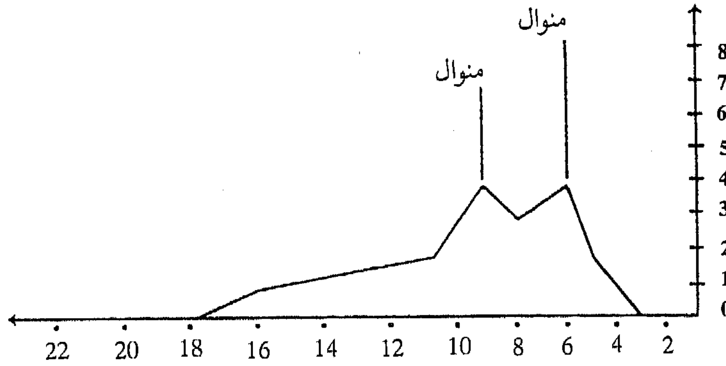
يتم حساب «المتوسط الحسابي»، بجمع قيم المرتبات، وقسمتها على العـدد الكلي للمرتبات، كالآتي:

4100	دولار
6000	
6000	4 قيـــم
6000	
8000	المتوسط العددي
9000	
10000	
11000	4 قيـــم
20000	
80100	دولار

$$\text{المتوسط الحسابي} = \bar{X} = \frac{80100}{9} = 8900 \text{ دولاراً}$$

أمـا «المتوسط العددي» فهو يساوي 8000، حيث تتساوى عدد القيم أعلاه وأسلفه كما هو موضح عاليه. قد تتعدد في التوزيع «قيم المنوال» [يطلق على هذا التوزيع متعدد قيم المنـوال]، وخـاصـة التـوزيعـات التي تحتوي على قيمتين من هذا النوع «قيمتان متكـررتـان للمتغير»، فهو أكثر شيوعاً. وفي المكتبات الكبرى حيث يكون عدد العاملين كبيراً، فقد نجـد أن عدداً كبيراً من العاملين يحصلون على مرتب يساوي 6000 دولار، وعـدداً كبـيراً آخـر من العـاملين يحصـل على أجر يساوي 9000 دولار. وفي مثل هذه الحالات ـ حيث تتعدد «قيم المنوال» ـ، يحدث تداخل فيها بينها، ويصعب على المرء،

التفرقة بين مجموعات القيم المختلفة، ويساعدنا أسلوب «المنوال» في تحديد متوسط أو «نزعة مركزية» للمرتبات، للمجموعات المختلفة للعاملين، التي قد تصنف كمهنيين، شبه مهنيين، ومبتدئين، كما هو موضح في شكل (4).



شكل (4): عرض مرتبات العاملين بالمضلع التكراري نموذج توزيع ثنائي

مقاييس التشتت Measures of Dispersion

تساعدنا «مقاييس التشتت» في التعرف على التباين من واقع البيانات المتعلقة «بالنزعة المركزية»، ولذا، كثيراً ما نستخدم واحداً أو أكثر من «المقاييس الموحدة»، التي تساعدنا على معرفة القيمة التلخيصية لأسلوب «النزعة المركزية» وأهميته.

يفيدنا معرفة أسلوب « التراوح : Rang » في التعرف على الفرق ما بين قيمة الحد الأدني والحد الأقصى لبيانات «التوزيع المتكرر» فهو يمثل (أي التراوح) القيمة المتوسطة لمجموع الحد الأدنى والأقصى للبيانات. فلو كانت بداية المرتب للكوادر المؤهلة 8000 دولاراً، والحد الأقصى لمرتباتهم يصل إلى 12500 دولاراً فإن «التراوح» في هذه الحالة يكون 4500 دولار (12500 – 8000 = 4500) دولار، وعادة يكتب «التراوح» بالارقام التي تعبر عن الحد الأدنى والأقصى كما يلي: 12500 دولار ـ 8000 دولار. ويستخدم «التراوح» للمقارنة بين المتغيرات، فمثلا، بداية المرتبات في الشمال الشرقي للبلاد يتراوح ما بين 6700 دولار ـ 9200 دولار، أما في الجنوب الشرقي، فيتراوح مابين 8000 دولار ـ 9400 دولار إذا «التراوح» يساوي 2500 دولار و 1400 دولار على التوالي.

وبالرغم من أن أسلوب «انحراف المتوسط الحسابي» لم يُعد شائع الاستخدام، إلا أنه لازال ذا فائدة لوصف التغير الذي يطرأ على «المتوسط الحسابي»، فهو يفيدنا في معرفة

مقدار «التراوح» (المتوسط الحسابي) للتغير الذي طرأ على المتوسط الحسابي ($\bar{X}$) لمجموعة من القيم. حيث يرمز لكل قيمة بالرمز X والانحراف عن المتوسط الحسابي = $\bar{X}-X$ ، ويرمز له بالرمز x ($x = \bar{X} - X$) ، ومعادلة الانحراف المطلق للمتوسط الحسابي = $\frac{\Sigma[\bar{X}-X]}{n}$ ويؤخذ مقدار القيمة للاختلافات بين كل قيمة ملاحظة وبين «المتوسط الحسابي» ثم يقسم n حاصل الطرح على عدد الملاحظات.

في جدول (7) ، يمثل العمود الرأسي المحتوي على $\bar{X}-X$ «أخذ قيمة مطلقة» ، ونعني بذلك تحويل أي فرق سلبي إلى موجب، والإبقاء على الفروقات الإيجابية كما هي [تحويل القيم السلبية إلى موجبة، والإبقاء على القيم الإيجابية كما هي].

تشير البيانات في جدول (7) ، أن «المتوسط الحسابي» للتداول الأسبوعي يساوي 1600 كتاباً، ويمكن الحصول على «انحراف المتوسط الحسابي بحساب فروقات «المتوسط الحسابي الأسبوعية» ثم جمع القيم المطلقة لهذه الفروقات، ثم قسمتها على عدد الأسابيع.

جدول (7) أرقام تداول المقتنيات في الفرع الرئيسي للمكتبة لكل أسبوع من الست أسابيع الماضية في العمود (X)

X	$\bar{X}-X$ أو x	القيمة المطلقة لـ x	
1200	400–	400	$\bar{X} = \frac{\Sigma X}{n} = \frac{9600}{6} = 1600$
1900	300	300	
1600	000	000	
1100	500–	500	
2000	400	400	$MD = \frac{\Sigma X}{n} = \frac{1800}{6} = 300$
1800	200	200	
$\Sigma X = 9600$		$\Sigma X = 1800$	
			حيث MD = مقياس التشتت

ويمكن تلخيص بيانات جدول (7) في التالي : $\bar{X} = 1600$, (MD) = 300 . تشير هذه الأرقام، أنه في خلال فترة 6 أسابيع، بلغ «معدل متوسط التداول» 1600 كتاباً، في حين بلغ معدل متوسط الانحراف، للمتوسط الحسابي 300 كتاباً أسبوعياً، بالزيادة أو النقص عن المعدل المتوسط العادي (± عن 1600 كتاباً) وذلك لمدة الستة أسابيع موضع الدراسة .

يُعد الانحراف المعياري ـ إلى حد ما ـ أكثر أهمية وأكثر صعوبة في تحديده وحسابه من «التراوح» أو من انحراف المتوسط الحسابي، ويفضل أسلوب الانحراف المعياري، لكونه يعتمد على خصائص رياضية ملائمة يفضلها علماء الرياضيات، ولذا نجده يرد بكثرة في الدراسات، ويرتبط بصفة أساسية بإجراءات الإحصاء الاستنباطي والاستنتاجي .

ولحساب الانحراف المعياري، علينا أخذ قيمة الانحراف عن المتوسط الحسابي (متوسط المتوسطات الحسابية)، ثم تربيع هذه القيم، وتجمع المقادير التي تنشأ عن هذه العملية، وتقسم على عدد الحالات الدراسية، ثم الحصول على «الجذر التربيعي» لها . ويُعد تربيع الانحرافات عن المتوسطات الحسابية ممارسة للمعايير الرياضية . أما في أسلوب «معدل الانحراف» فمطلوب منا تجاهل العلامات (علامات التربيع)، ولكن يبدو أن هذه الإجراءات غير مرغوب فيها ـ إلى حد ما ـ من وجهة النظر الرياضية . فعندما نقوم بتربيع الانحرافات، نعمل في النهاية على التخلص من العلامات، وننتهي بالحصول على رقم مطلق للانحراف عن المتوسط الحسابي (بصرف النظر عن وجود علامات) . إضافة إلى أن تربيع الانحرافات يضفي على العمليات الرياضية التي نقوم بها وزناً، كما أنه يمدنا بقاعدة بيانات رياضية، يمكن الاعتماد عليها في دراسة معدلات أكثر للانحرافات لمجتمع الدراسة موضع البحث، وبما أننا في النهاية نأخذ بالجذر التربيعي، فنحن نعود مرة أخرى إلى الأرقام الحقيقية التي تخص عينة البحث موضع الدراسة . والمعادلة العامة المستخدمة لحساب الانحراف المعياري هي :

$$S = \sqrt{\frac{\Sigma (X - \bar{X})^2}{n}}$$

وإذا استخدمنا البيانات التي وردت في المثال السابق، يمكن حساب الانحراف المعياري بالطريقة التالية :

التداول الأسبوعي للمواد المكتبيـة	الانحراف عن المتوسط الحسابي لمدة 6 أسابيع	تربيع الانحراف
X	$(X-\bar{X})$	$(X-\bar{X})^2$
1,200	400–	160,000
1,900	300	90,000
1,600	000	0,000
1,100	500	250,000
2,000	400	160,000
1,800	200	40,000
		700,000

الانحراف المعياري $= S = \sqrt{\frac{700,000}{6}} = \sqrt{116,666,67} = 341,57$

حيث :

$n = 6$

$\bar{X} = 1,600$

$\Sigma (X-\bar{X}) = 700,000$

وتفسير الانحراف المعياري ، يكون سهلاً لو كان للمتوسط الحسابي شكل التوزيع الطبيعي (التوزيع الطبيعي سينـاقش بعـد قليل) ويكفي ـ حاليا ـ أن نقـول ، أن الانحراف المعياري للمتوسط الحسابي (زائد وناقص واحد) (±)، ينطبق على ثلثي حالات البحث في حالـة التـوزيع الطبيعي أوشبه الطبيعي ، وفي المثال المطروح ، فإن الانحـراف الميعاري 1800 ± 342 ينطبق على 4 حالات من 6 حالات ، أي ما يعادل ثلثي الحـالات .

إذا قمنا بالتركيز على البيانات الخاصة بالمكتبات العامة في الجدول (8) ، الذي يصف عدد الكوادر المؤهلة العاملة بمختلف أنواع المكتبات في الجزء الجنوبي الشرقي للبلاد ، لاستطعنا تلخيص بياناتها (أي المكتبات العامة) كالآتي :

الحالة	عدد الكوادر المهنية
الباما	119
فلوريدا	205
جورجيا	238
كنتكي	63
ميسيسبي	112
شمال كارولينا	127
جنوب كارولينا	85
تينيسي	177
فيرجينيا	261
المجموع	1387

$\bar{X} = \frac{1387}{9} = 154$ (عوضت إلى أقرب عدد صحيح)

وبحساب الانحراف المعياري نجده يساوي: 65

$\therefore \bar{X} \pm S = 154 \pm 65$

$= 89 - 219$

وقد كان من المتوقع إن نجد ـ وقد وجدنا بالفعل ـ الانحراف المعياري للمتوسط الحسابي ينطبق على (5) ولايات من (9) وهو رقم يقترب بالفعل من ثلثي الحالات، وبالتالي من التوزيع الطبيعي.

وبما أن الانحراف المعياري يُعد مقياساً معيارياً للتشتت، فإنه يمكن استخدامه للمقارنة بين تكافؤ مجموعتين أو أكثر أو عدم تكافؤها، ولو كانت المجموعات قابلة للمقارنة، ووجد إختلاف كبير في الانحرافات المعيارية فيما بينها، فإن هذا يؤكد إحتمال عدم وجود تكافؤ بينها. وعادة، لو تمت المقارنة بالاعتماد ـ فقط ـ على الانحرافات المعيارية، فمن الممكن أن يقودنا ذلك إلى تفسير خاطىء لنتائج الدراسة. ولتجنب هذا الخطأ، علينا اتباع منهجية تعتمد على ما يسمى «بمعامل الانحراف».

معامل الانحراف Coefficient of Variations

قد تختلف قيم الانحرافات المعيارية في بعض الأحيان ـ بصورة كبيرة ـ ولكن ليس بالضرورة أن يعكس ذلك حقيقة تجانس المجموعات غير المتكافئة، ومثال لذلك:

إذا كان الانحراف المعياري لميزانيات فرع مكتبة عامة يقدر بـ 1000 دولار في عام 1970, 1300 دولار في عام 1975 هذا لا يعني أن عدم التكافؤ المالي بين أفرع المكتبة يتدهور إلى الأسوأ خلال الخمس سنوات التي غطتها الدراسة. إن المقارنة بين الفترتين (1975, 1970)، يجب أن تستند على معرفة نسبة الانحراف المعياري للمتوسط الحسابي أي

تستنـد على ما يسمى «بمعـامـل الانحـراف»، فإذا حصلنـا على المتـوسط الحسـابي للميزانيات لعـام 1970 ,1975 نستـطيع أن نحسب هذا المعامل بسهولة، وذلك باتباع المعادلة التالية :

$$V = \frac{S}{X}$$

حيث: V = معامل الانحراف، S = الانحراف المعياري، و $\bar{x}$ = المتوسط الحسابي

1970: المتوسط الحسابي لميزانية فرع المكتبة	$= \bar{X} = 10000$	دولار
الانحراف المعياري	$= S = 1000$	دولار
معامل الانحراف	$= V = \frac{1000}{1000} =$,100	
1975: المتوسط الحسابي لميزانية فرع المكتبة	$= \bar{X} = 14000$	دولار
الانحراف المعياري	$= S = 1300$	دولار
معامل الانحراف	$= V = \frac{1300}{14000} =$,093	

يمكننـا أن نستنتـج من ذلك أن عدم التكافؤ في ميزانيات فروع المكتبة، انخفض فعلياً خلال الأعـوام 1975,1970 بينـما ارتفع الانحراف المعياري عن نفس الفترة، أما معدل متوسط الميزانية، فقد ارتفع نسبياً، وبدقة أكثر، فإن معدل التغير يساوي 093,) (100, – / أي = 07, أو 7% ، انخفاض في عدم التكافؤ (أي اقتراب من التكافؤ).

التوزيـع الطبيعـي The Normal Distribution

يُعد الإحصاء فرعاً من فروع الرياضيات، ولذا يجب علينا دراسة وتعلم بعض فروع التـوزيع الـرياضي، لنتمكن من استخـدام وفهم المنـاهـج الإحصـائية، وأول فروع التـوزيعـات الـرياضية التي سنقـوم بدراستها، هو مايسمى «بالتوزيع الطبيعي»، أو «المنحنى الطبيعي». والشكل رقم (5) يمثل أحد أشكال التوزيع الطبيعي. والملاحظ أن كل أشكال التوزيع الطبيعي تتشابه في الخصائص، أما الشيء الوحيد الذي قد لا يتم التشابه فيه، هو المتوسط الحسابي أو الانحراف المعياري للتوزيع. والنقاط الأساسية التي يتبعها المنحنى في شكل (5) ، مأخوذة من معادلة رياضية (لن يتم مناقشتها الآن).

حقيقة الأمر، أن جميع أشكال «المنحنيات الطبيعية» مؤسسة على هذه المعادلة ومبنية

جدول (8) عدد العاملين المؤهلين، مصنفين حسب نوع المكتبة المنتمين إليها، والولاية، والإقليم

الولايــة	مكتبات أكاديمية	مكتبات التعليم	مكتبات عامة	مكتبات مدرسية/أطفال	مكتبات متخصصة	مكتبات * الولاية	المجموع
الباما	125	17	119	151	30	10	452
فلوريدا	305	10	205	424	28	16	988
جورجيا	277	21	238	952	27	17	1532
كنتكي	102	11	63	668	18	14	876
ميسيسبي	118	10	112	118	13	4	375
شمال كارولينا	307	24	127	395	30	20	903
جنوب كارولينا	140	8	85	225	29	15	502
تينيسي	199	15	177	260	38	17	696
فيرجينيا	341	3	261	742	44	41	1432
المجموع	1914	119	1387	3935	257	144	7756

* تتضمن مكتبات مؤسسات الولاية، مكتبات الولاية، مراقبي المكتبات المدرسية التابعين للولاية.

مصـدر المعلومـات: ماري أدنا أنديرس Mary Edna Andres ، جداول مسح ميداني (أتلنتا: معهد جورجيا التكنولوجي 1975 ، ص 80.

عليهـا. أما العامل (أو العوامل) الذي يفرق بين معادلة وأخرى. عند تطبيقها لرسم المنحنيات البيانية، هي الأرقام المتعلقة بالمتوسط الحسابي أو/ و الانحراف المعياري.

تعتمـد كل التـوزيعات التي سنتعامل معها في هذه الدراسة، على معادلة مناظرة، نستنـد إليهـا في تحديد مواقـع النقـاط التي تشكـل «المنحنى الـطبيعي»، وتُعـد هذه التوزيعات، صورة طبق الأصل من «توزيعات التكرار»، وإن اختلفت عنها في اتسامها بالاستمـرارية والإنسيابية، نتيجـة لتشكلهـا من أعداد لا نهائية من النقاط، كذلك، يعتمـد «المنحنى الـطبيعي» في تكـوينـه على معـادلة رياضية، بينما تعتمد «توزيعات التكرار» على البيانات النابعة من التجارب العملية والملاحظة، وعندما تكون قيمة المتغير مرسومة بيانياً على المضلع التكراري، اعتماداً على مدى تكرارها، فإن هذا يعني أن قيمة المتغير تشكلت بيانياً، بناء على احتمالية مؤكدة في التوزيع الاحتمالي الرياضي.

وتتميز مجمـوعة «التوزيعات الطبيعية»، بخصائص، تُعد مفيدة فيما يتعلق بتطبيق الانحـراف المعياري، وهـو الأمـر الـذي سنقوم بتوضيحه في الفصل الخاص بدراسة

الإحصاء الاستنباطي، أما هنا فسنكتفي بذكر هذه الخصائص :-

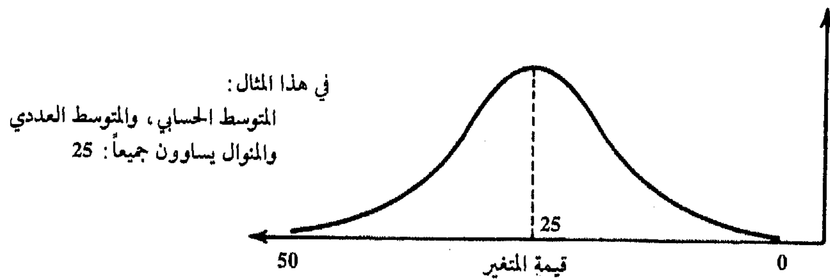
1 ـ يُعد المتوسط الحسابي، والمتوسط العددي، والمنوال، متساوية في القيمة، فإذا قمنا برسم خط رأسي مستقيم يبدأ من أعلى نقطة في المنحنى، سنجد لدينا خطاً يقسم المنحنى إلى نصفين متساويين، ونقطة التقاء هذا المستقيم مع الخط الأفقي (الاحداثي الأفقي) هي التي تحدد قيمة القياسات الثلاث (المتوسط الحسابي، المتوسط العددي، المنوال)، وتمثل «النزعة المركزية»، للتوزيع.

2 ـ يتصف التوزيع بالتماثل، فالمساحات على جانبي الخط المتقطع متساوية. وهذا يعني أن التوزيع العددي للحالات في كلا الجانبي مساو للآخر.

3 ـ النهايات الطرفية للمنحنى (الذيول)، تمتد إلى مالا نهاية، بدليل أنها لا تلامس الاحداثي الأفقي، فهي متقاربة وموازية له.

4 ـ يمثل هذا المنحنى عدداً غير محدود أو ـ بالتأكيد ـ عدداً كبيراً من الحالات (أنظر شكل 5).

يكتسب «المنحنى الطبيعي «صفة الشيوع والانتشار، لأن أحداث الحياة، هي أقرب إلى نموذج التوزيع الطبيعي. ومن الأشياء التي تدلنا على أن هذا التوزيع طبيعي، وجود غالبية الحالات بالقرب من مركز المنحنى، فلو قمنا برسم خطوط متعامدة مع قاعدة المنحنى تبدأ من الانحراف المعياري ± ، في كل من جانبي المتوسط الحسابي، سنجد المنطقة الواقعة تحت المنحنى مابين المتوسط الحسابي وكلا الخطين تمثل 34% من الحالات في كل من الجانبين، لذا فمفهوم «المعيار» (وحدة انحراف معياري واحدة من المتوسط الحسابي) يتضمن 34% من الحالات في التوزيع الطبيعي أو شبه الطبيعي، بصرف النظر عن حجم التوزيع.

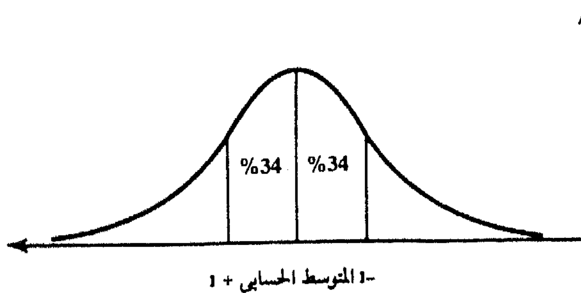


شكل (5): التوزيع الطبيعي، مع المتوسط الحسابي لمجتمع دراسة يساوي 25.

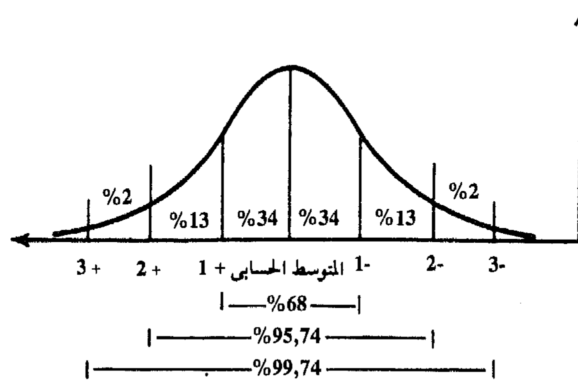
وقد تم تحديد موقع «وحدة الانحراف المعياري» على كل من جانبي المتوسط الحسابي في شكل (6) ، ونجد أن «وحدة الانحراف المعياري» التي تقع على يسار المتوسط الحسابي تتضمن 34% من الحالات التي قيمتها أقل من المتوسط الحسابي، أما الوحدة التي تقع على يمين المتوسط الحسابي فهي تتضمن 34% من الحالات التي قيمتها أعلى من المتوسط الحسابي، ومجموع الحالات في الجانبين يمثل 68% ، أو بما يساوي ثلثي المجموع الكلي للحالات.

بما أن «وحـدة الانحراف» (المسافة الواقعة على الخط القاعدي للمتوسط الحسابي) ثابتة، فبإمكاننا التحرك أسفل المنحنى على هذا «الخط القاعدي» بنفس قيمة «وحدة الإنحراف» فنحصل على مسافات متماثلة وثابتة لكل وحدة. وعند تطبيق ذلك المنظور على شكـل (7) ، سنـلاحظ، أن الأجـزاء التي تنشأ تأخذ في الصغر كلما ابتعدنا عن مركز المنحنى، وذلـك بسبب الشكـل المنحـدر للمنحنى، ويشـير شكل (7) أن الانحراف المعياري الثـاني أضاف 13% (لكل جانب)، والانحراف المعياري الثالث أضاف 2% على كل من الجانبين.

وغالباً ما ينصب اهتمامنا على قيمة وحدة أو وحدات الانحراف المعياري الواقعة على جانبي المتـوسط الحسـابي، وفي هذا الصـدد، فإن الوحدة الأولى للانحراف المعياري للمتوسط الحسابي تحتوي على 68% من الحالات (34% + 34%) + (± 1) ، والتي يشار إليهـا ـ تجاوزاً ـ «بثلثي الحـالات»، الـوحـدتـين الأولي والثـانية للانحـراف المعياري للمتوسط الحسابي تحتويان على 95% من الحالات (34% + 34% + 13% + 13% + (1 ±)، والمجمـوع الكـلي لوحـدات الانحـراف المعياري الثـلاث، الواقعة على جانبي المتـوسط الحسـابي تسـاوي أقـل قليلاً من 100% من العـدد الكي للحالات. ويمثل



شكل (6) توزيع طبيعي، يشير إلى وحدة انحراف معياري واحدة أسفل وأعلى المتوسط الحسابي.



شكل (7) توزيع طبيعي يشير إلى مواقع الوحدات 3,2,1 للانحراف المعياري، لكل جانب من جانبي المتوسط الحسابي

الانحراف المعياري أهمية خاصة للإحصاء الاستنباطي والاستنتاجي فعندما يستخدم لتلخيص مجموعة من البيانات، فهو يساعد على توصيف التشتت بطريقة واضحة. ونستنتج من العرض السابق أن الانحراف المطلق «للمتوسط الحسابي» غالباً ما يكون مفيداً في أغراض الإحصاء الوصفي المحض. كما أن الانحراف المعياري، يمكن أن يفيد عندما تكون البيانات متعلقة «بالتوزيع الطبيعي».

على أي حال، فإن «الانحراف المعياري»، هو إجراء يمكن اتباعه للتحليل الوصفي، عندما لا يكون هناك حاجة ملحة إلى تحليل استنتاجي أو استنباطي، وذلك لأن «الانحراف المعياري» لا يتطلب إجراءات رياضية معقدة كما أنه يتميز بالمرونة، ويضفي على التحليل مظهراً علمياً جذاباً.

الدرجة المعيارية، Z Scores:

(Z) أو الدرجة المعيارية، مستنبطة بصفة مباشرة من «الانحراف المعياري»، ويمكن استخدامها مع أي نمط من أنماط التوزيع، ولكنها تكون أكثر فعالية عندما يأخذ التوزيع الشكل الطبيعي، وهذا ما يحدث عندما يتساوى «المتوسط الحسابي» مع «المتوسط العددي»، وتكون قيمة الوحدة الأولى من الانحراف المعياري على جانب المتوسط الحسابي مساوية لثلثي عدد الحالات، وتقدر قيمة الدرجة المعيارية (Z)، بناء على عدد الوحدات للانحراف المعياري، فهي تشير إلى عدد وحدات «الانحراف المعياري» أسفل وأعلى المتوسط الحسابي» ونعني أن قيمتها تساوي عدد النقاط التي تمثل الحالات (أو الحالات نفسها).

وقد اكتسب كثير من الباحثين خبرتهم عن الدرجة المعيارية (Z)، عندما وجدوا أنفسهم في موقف المتعامل معها، وخير مثال على ذلك، اختبارات الذكاء التي شاع استخدامها، وساعدتنا على فهم واستخدام توزيع الدرجات المعيارية (IQ)، كتوزيع معياري طبيعي. فمثلاً، إذا افترضنا وجود درجات معيارية (IQ)، حيث المتوسط الحسابي يساوي 100، والانحراف المعياري 15، إذا حسب القاعدة 68% (ثلثي الحالات)، من أفراد مجتمع الدراسة (الحالات)، يحصلون على مابين 115,85، فإذا كان هنالك شخص حاصل على 130 درجة، فهو أعلى من المتوسط الحسابي، بوحدتي انحراف معياري، وهذا يضعه في مصاف الـ 2% من مجتمع الدراسة ممن حصلوا على أعلى الدرجات.

والمثال التالي يوضح لنا كيفية تطبيق هذه المبادىء في مجال المكتبات والمعلومات، وهو يتعلق بمعضلة عينة الدرجة المعيارية (Z). فقد ورد في دراسة مسحية لمرتبات الإداريين العاملين بالمكتبات، أن المتوسط الحسابي للمرتبات بلغ 17940 دولاراً، والانحراف المعياري 4960 دولاراً، وأراد أحد الإداريين ممن بلغ راتبهم 25000 دولاراً، أن يعرف موقفه من هذا التوزيع، وهو على قناعة بأن مرتبات الإداريين العاملين بالمكتبات موزعة طبيعياً (أي تتبع نظام التوزيع الطبيعي)، ولمعرفة حساب موقعة النسبي (أي موقعة النسبي من التوزيع) نطبق عليه المعادلة التالية:ـ

$$Z = \frac{(\bar{X}-X)}{S} = \frac{17940-25000}{4960} = \frac{7060}{4960} = 1{,}42 \text{ دولاراً}$$

يمثل الرقم 1,42 عدد وحدات الانحراف المعياري، التي تبعد عن المتوسط الحسابي، حيث يقع مبلغ الـ 25000 دولاراً، وبما أنها أكبر من المتوسط الحسابي (حيث بلغ المتوسط الحسابي 17940 دولاراً)، فنحن نعلم أنها سوف تقع بهذا القدر من الوحدات (1,42) أعلى الـ 50% من الحالات، وباستخدام جدول «مناطق تحت المنحنى الطبيعي» (الملحق: جدول رقم (2))، نستطيع أن نحدد قيمة النسبة المئوية لـ 1,42 وحدة انحراف معياري.

ولكي نستخدم الجدول الموجود في الملحق، نبدأ بقراءة العمود الأول (XQ) حتى نصل إلى الرقم 1,42، ثم نتحرك مع الأرقام أفقيا حتى نصل إلى العمود (,02) سنجد الرقم ,4222، وبتحويل هذا الرقم إلى نسبة مئوية سيصبح 42,22% (أو 42% تجاوزاً). وهذا يعني أن مرتب هذا الإداري أكبر من 42% من المرتبات التي تعلو المتوسط الحسابي

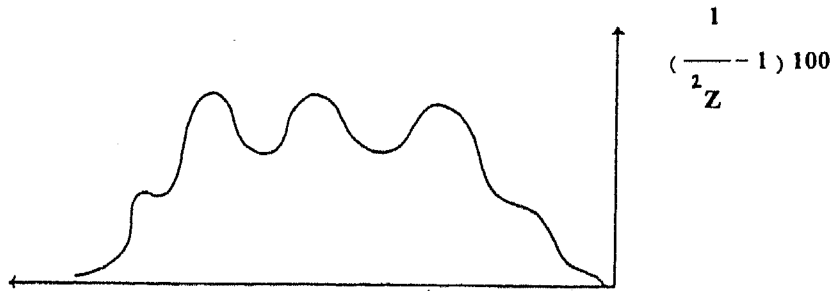
للمرتبات، وبعبارة أخرى فإن مرتب هذا الإداري أكبر من مرتبات مجموع الإداريين (حيث أن المتوسط الحسابي لمرتب هذا المدير يقع فوق المتوسط الحسابي لمجموع المرتبات بـ 42 درجة، فإذا ما جُمعت الدرجات أسفل المتوسط الحسابي مع درجاته أعلى المتوسط الحسابي، يكون مجموعة كالآتي:

(50 + 42 = 92). نستنتج من ذلك، أن 92% من الإداريين يحصلون على مرتبات أقل، و 8% من الإداريين يحصلون على مرتبات أعلى، وتبدأ مرتبات هذه الفئة الأخيرة من 25000 دولاراً فأكثر (مع ملاحظة أننا افترضنا توزيع مرتبات الإداريين ـ التي خضعت للبحث ـ توزيعاً طبيعياً).

نظريــة تشيبيشيف Chebyshev's Theorem

طبقاً لنظرية تشيبيشيف، يمكن استخدام «الانحراف المعياري» لتحديد المناطق غير الموزعة طبيعياً تحت المنحنى، إذا يكون لدينا ـ في بعض الأحيان ـ أسبابنا لنعتقد بأن التوزيع الذي نتعامل معه ينتمي إلى نمط «التوزيعات المتعددة»، وفي هذه الحالة لن يفيد تطبيق «الانحـراف المعياري»، حيث أن المنحنى لن يكون «منحنى طبيعياً» وقد نتعامل مع نمط توزيع يتشابه مع ماهو مبين في شكل (8)، وفي حالة كهذه، ومع مثل هذا المنحنى، نستطيع تطبيق المعادلة التالية

$$\left(\frac{1}{Z^2} - 1\right) 100$$



شكل (8): توزيع متعدد النماذج

إذا علمنا من المثال السابق، الخاص بمرتبات الإداريين أن توزيع هذه المرتبات يأخذ نمط التـوزيعـات «متعددة النماذج»، فإننا نلجأ إلى استخدام المعادلة السابقة لتعديل التفسير المتعلق بالدرجة المعيارية (Z) وتكون الصيغة الجديدة، لتعويض المعادلة كالآتي:

$$\%51 = (,51)100 = (,495-1)100 = \left(\frac{1}{2,02}-1\right)100 = \left(\frac{1}{(1,42)^2}-1\right)100$$

ونحصل من هذه الصيغة الجديدة، على تقدير أكثر تحفظاً، للموقع النسبي لمرتب هذا الإداري، فبدلاً من وضعه ضمن الـ 8% من المرتبات الأكثر (في حالة التوزيع الطبيعي)، نفترض وضعه ـ هنا ـ ضمن الـ 49% من المرتبات الأكبر قيمة.

تمرينات الفصل الأول:

1. بالإشارة إلى جدول 8 (عدد الكوادر المؤهلة المصنفة بناء على نوع المكتبة التي يعملون بها، والولاية، والإقليم).
   أ ـ ماهي النسبة المئوية للمكتبيين، العاملين بالمكتبات العامة، بولاية نورث كارولينا؟
   ب ـ ماهي المتوسطات الحسابية، المدى، والانحراف المعياري لتوزيع المكتبات العامة في الجنوب الشرقي؟
2. بالإشارة إلى جدول 6، الذي يتضمن توزيع التكرار لعدد المحطات لسبعة عشر مجلس محلي.
   أ ـ إحسب المتوسط الحسابي، المتوسط العددي، منوال التوزيع (تذكر أن هناك سبع عشرة حالة بحثية).
   ب ـ حدود الانحراف المعياري.
   ج ـ أوجد الدرجة المعيارية (Z) التي تطابق قيمة 16 للمتغير.
   د ـ إرسم مدرج التكرار للبيانات.

المراجع والقراءات:

Babbie, Earl R. *The Parzctice of Social Research.* Belmont, Calif.: Wadsworth, 1975.

Blalock, Hubert M., Jr. *Social Statistics.* 2d ed. New York: McGraw-Hill, 1972, p. 11-88.

Fields, Caraig. *About Computers.* **Cambridge, Mass.: Winthrop, 1973. Glass, Gene, and Julian C. Stanley.** *Statistical Methods for Education and Psychology.* Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Hoadley, Irene, and Helen Clarke, eds. *Qunatitative Mehtods in Librarian-ship: Standards, Research, Management.* Westport, Conn.: Greenwood, 1972.

Huff, Darrell. *How to Lie with Statistics. New York: Norton, 1954. Kerlinger, Fred N. Foundations of Behavioral Research. 2d ed. New York: Holt, Rinehart and Winston, 1973.*

Loether, Herman, and Donald McTavish. *Descriptive Statistics for Sociologists. Boston: Allyn and Bacon, 1974.*

Sanders, Virginia. ·Measurement and Statistics. New York: Oxford University Press, 1958.

Siegel, Sideney. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill, 1956.

الفصل الثاني

المعاينة الاحصائية

SAMPLING

الفصل الثاني

المعاينة الاحصائية [1]

SAMPLING

تعاملنا في الفصل الأول مع الإحصاء الوصفي، وهو أحد الطرق التي تعيننا على تلخيص الخصائص لمجموعة مجالات بحثية على ضوء متغيرات محددة، ومن ثم تحويلها إلى أرقام قابلة للتفسير (مثال: المتوسط الحسابي و / أو الانحراف المعياري). وقد نشأ مصطلح «الإحصاء الوصفي»، لأننا نصنف مجموعة حالات البحث التي بين أيدينا عن طريق البيانات التي قمنا بجمعها عن هذه الحالات. وفي بعض البحوث من الممكن أن نخرج بمؤشرات عن مجموعة كبيرة من الحالات البحثية، استناداً على معلومات تجمعت لدينا عن مجموعات فرعية صغيرة تتضمنها تلك المجموعة الكبيرة، وفي مثل هذه الحالة، يطلق على تلك المجموعات الصغيرة اسم «العينة». واختيار هذه العينات يتم بناء على قواعد ما يسمى بـ «نظرية المعاينة الإحصائية».

تعتمد اختبارات الإحصاءات الرسمية جميعها على فرضية عشوائية الاختيار للعينة موضع الدراسة، ولذا فإن المعاينة الإحصائية والنظرية الخاصة بها، يلعبان دوراً في غاية الأهمية، فيما يخص الإحصاء الاستنباطي أو الإستنتاجي. وسوف نتناول في هذا الفصل، شرح المفاهيم الخاصة بالعشوائية، بجانب شرح بعض المصطلحات الأخرى المتعلقة بالمعاينة الإحصائية. ونأمل في هذا الصدد أن نزود القارىء بمزيد من الإيضاحات حول الاعتبارات الرئيسية الخاص بتحديد وتأطير خطط المعاينة الإحصائية.

وقد تطورت تقنيات المعاينة الإحصائية إلى حد كبير، حتى أن البحوث أصبحت تعتمد إلى حد بعيد على المعاينات الإحصائية، وأصبح الباحثون يلجأون إلى استشارة

---

(1) ترجم أيضاً بمصطلح «المعاينة»، وقد فضلنا استخدام مصطلح «المعاينة الإحصائية» لرأينا أنه أقرب إلى المفهوم والسياق الموضوعي (المترجمان).

الخبراء في هذا المجال. ومن المهم بمكان أن يتزود المكتبيون ـ كغيرهم من الباحثين ـ بالمعلومـات الجوهرية اللازمة لكيفية استخراج العينات، وعليهم في ذلك الاستعانة بخبراء المعاينة الإحصائية. ومن ناحية أخرى، يمكن في كثير من الحالات أن يقوم باحث منفرد أو مجموعة من الباحثين في مجال المكتبات، بتطوير خطة جيدة للمعاينة الإحصائية بدون الاستعانة برأي الخبراء، وسيجد القارىء من خلال التوضيح التالي، المعلومـات الأساسية المطلوبة لفهم المشكلات الرئيسية، والخطوط العريضة اللازمة بإجراءات المعاينة الإحصائية. ويتضمن جدول رقم (9) الإطار العام لأنواع العينات التي سوف تناقش في هذا الفصل. والغرض الرئيسي لعملية المعاينة الإحصائية، هو تزويدنـا بالمعلومـات التي تعيننا على تعميم النتائج على مجتمع الدراسة Population والذي أخذنا منه العينة.

ومجتمع الدراسة أو المجتمع المبحوث عبارة عن مجموعة من الأفراد أو الأشياء التي تمتاز بسمات أو خصائص معينة نود قياسها. فمثلاً لو كان اهتمامنا هو بالمستفيدين من المكتبة، فإن مجتمع الدراسة المعني هنا هو المستفيدين من المكتبة.

ولا يُعـد استخـدام أسلوب العينة في بعض الحالات البحثية ضرورياً، بل على العكس، قد يكون في استخدام الإحصاءات الكلية لمجتمع الدراسة مميزات أكثر، مثال لذلك: إذا أراد الباحث أي يُجري بحثه على مجتمع عدد أفراده صغير نسبياً، فإن استخدام العدد الكلي لذلك المجتمع في هذه الحالة أفضل بكثير من استخدام أسلوب العينة، ففي حالة مكتبة أكاديمية (مكتبة كلية)، تقوم بخدمة عدد من الطلبة الجامعيين

جدول (9) تقنيات المعاينة الإحصائية التي سيتم مناقشتها في هذا الفصل.

الإحتمالية	غير الإحتمالية
عينة عشوائية *	عينة غرضيــة
عينة منتظمـة *	عينة حصصية
عينة طبقيــة *	عينة الصدفـة
نسبيـة	
غير نسبية	

* يشير إلى إمكانية استخدام البيانات عن طريق تقنيات الاستنتاج المعتمدة (الرسمية).

يقدر عددهم بـ 900 طالباً، أو في حالة مكتبة متخصصة تقوم بخدمة عدد محدود من العاملين، من الأفضل أن نقوم بدراسة مجتمع الدراسة بالكامل، إذ أن تكلفة البحث في هذه الحالة لن تكون كبيرة، في حين أننا سنحصل على نتائج مؤكدة وموثوق بها. وفي حالة، إذا ما أراد باحث ما أن يقوم بحساب «متوسط سعر» الكتاب، لمجموعة كتب يقدر عددها بـ 100 كتاب فيمكنه ببساطة حساب هذا المتوسط عن طريق العدد الكلي للكتب، بدون اللجـوء إلى أسلوب العينـة، وبـذلـك يلقى استخـدام أسلوب الإحصاءات الكاملة استحساناً من الجمهور، سواء فيما يخص استجابتهم للرد على الاستفسارات أو قبولهم لنتائج البحث. ومن البديهي أن نتائج التحليل التي اعتمدت على إحصاءات كاملة تكون أكثر دقة، ويسهل على الباحث في هذه الحالة الدفاع عن أسلوب جمعه للبيانـات أو عن النتـائـج التي توصل إليها، إضافة إلى أن استخدام الإحصاءات بكاملها يُغني عن الاستعانة برأي خبراء المعاينة الإحصائية.

ولكن على صعيد الواقع لا يملك الباحثون الوقت والمال اللازمين، لتحليل مجتمع الدراسة بالكامل. ومن البديهي أنهم يرغبون في عمل أقصى جهد في حدود الإمكانات المتاحة والمحدودة التي يملكونها للتوصل إلى بعض النتائج حول مجتمع الدراسة بكامله، وهنا تبرز أهمية أسلوب المعاينة الإحصائية.

وفيما يخص دقة النتائج، فالعينة التي تؤخذ بطريقة علمية صحيحة، يمكن أن تعطينا نتـائـج تقـترب في جودتها من النتائج التي نحصل عليها من استخدامنا للإحصاءات الكاملة، بالإضافة إلى أنه مهما بلغت ضخامة إمكانات البحث سواء من زاوية الوقت المتاح أو الاعتمادات المالية، فلن نستطيع بأي حال من الأحوال تغطية نفقات الخبراء أو الخـدمـات المصـاحبة، للأخذ بعينة بحثية كبيرة الحجم، ناهيك عن تغطية النفقات اللازمة لتغطية المجتمع الدراسي بالكامل.

المعاينة الإحصائية الاحتمالية Probability Sampling

تُصنف العينات إلى فئتين رئيسيتين: العينات الاحتمالية، والعينات غير الاحتمالية. تعتمـد نظرية اختيار الفـروض الإحصـائية على الفـرضية القائلة بأن العينة الخاضعة للبحث تتكون من وحدات أو مفردات يتميز كل منها باحتمالية أن تكون ممثلة في العينة المختارة من مجتمع الدراسة، وبتعبير آخر، هي «عينة محتملة الظهور».

يجب إجـراء المعـاينة الإحصائية الاحتمالية، من خلال أسلوب المعاينة العشوائية. ويجب أن يؤخـذ في الاعتبـار مبـدأ احتمالات الظهور، عند القيام بعملية تقدير حجم

العينة. وقد تم إجراء العديد من الدراسات حول الأخطاء المعيارية، وانحرافات خطط اختيار العينات وتميزها على ضوء تقنيات نظرية الاحتمالات، وتبين أن اختيارنا لأفضل خطط المعـاينة الإحصائية الملائمة لاحتياجاتنا، يتوقف على مدى درايتنا ومعرفتنا بهذه التقنيات، وتعيننا المعاينة الإحصائية الاحتمالية على حساب الخطأ المعياري في تقديراتنا وتعيننـا على معـرفة حدود الثقة (ستناقش فيها بعد) للقيمة الحقيقية لمجتمع الدراسة، وذلك من خلال تحليل عينات البيانات التي قمنا بجمعها. وبذلك، نستطيع التأكد من صحة ودقة التقديرات التي توصلنا إليها، وذلك عكس أسلوب المعاينة الإحصائية غير الاحتمالية.

والغـرض من إجـراء عملية المعاينة الإحصائية، هو قياس بعض خصائص العينة الخـاضعـة للبحث، من أجل تقدير طبيعة هذه الخصائص في مجتمع الدراسة المعني. وحتى تكون تقديراتنا صحيحة، يجب أن تكون العينة المختارة مُمثلة ـ بقدر الإمكان ـ لمجتمع الدراسة، ويُعد أسلوب العينة الاحتمالية هو الأفضل لحصولنا على عينة أقرب ما تكون لتمثل مجتمع الدراسة.

المعاينة العشوائية البسيطة Simple Random Sampling

تُعد العينة العشوائية البسيطة، أحد الأنواع الرئيسية للمعاينات الإحصائية الاحتمالية. وفي هذا النوع من العينات، تتاح الفرصة لكل فرد من أفراد مجتمع الدراسة أن يكون ضمن أفراد العينة المختارة. والباحث ـ في هذه الحالة ـ لا يعتمد على تقديره الشخصي في ضم أو استبعـاد فرد أو مجمـوعـة أفراد من مجتمع الدراسة عند القيام بإجراءات اختيار العينة. وبالرغم من أن شرح إجراءات العينات العشوائية وتوضيحها يعتمد على الأسلوب الرياضي المعقد والغامض إلى حد ما إلا أن فهم طبيعة العينات العشوائية، يُعد بسيطاً للغاية. ولا يتطلب دراية كبيرة بالمعطيات الرياضية، ولذا فمن السهل تطبيقها، والفضل في ذلك يرجع إلى من قام بالأعداد المسبق لجداول الأرقام العشوائية التي تستخدم في هذا المجال.

دعنا نفترض أننا نود استخراج عينة تقدر بـ 80 طالباً جامعياً من بين قائمة تتكون من 480 طالب، سيمثلون مجتمـع الـدراسـة المـزمع بحثه، كمستفيدين من مكتبة تختص بالعلوم والتكنولوجيا. وما علينا في هذه الحالة، إلا ترقيم الأسماء تصاعدياً، بداية برقم (1) وحتى الرقم (480) ثم مراجعة أي جدول للأرقام العشوائية (أنظر الملحق: جدول رقم (1)، نختار أي صف مكون من ثلاث أرقام (أو أي عدد من الصفوف يمكن أن تعطي ثلاثة أرقام)، ونبدأ من أي نقطة عشوائية. وينبغي إلا نعتمد اختيار نقطة البداية

برقم يؤدي إلى تضمين بعض الحالات (الأسماء) التي نفضلها شخصياً، وإلا اعتبر ذلك تدخـلاً في الإختيار العشوائي، وبالتالي يؤثر على نتائج البحث. بعد ذلك نبدأ بقراءة الأعداد المكونة من ثلاثة أرقام ـ بمسار أفقي أو رأسي ـ ونستخرج كل الأرقام التي تقع ما بين 1 و 480 حتى نصل إلى عدد الحالات المطلوبة (80 حالة) وعندما نقع على رقم أعلى من رقم 480 فعلينا إهماله، لأنه لا يطابق أي من الحالات البحثية لدينا، كذلك إذا تكرر ظهور رقم أكثر من مرة علينا إهماله بعد الأخذ به عند ظهوره أول مرة. ونظراً لسهولة هذا الأسلوب وبساطته، فهو يبدو غير قابل للتصديق بل ومحيراً للغاية.

وكمثال تطبيقي لما سبق شرحه، لنفترض أننا بصدد اختيار عينة مكونة من 5 أفراد، من بين مجتمع دراسة يتألف من 50 شخصاً، ينبغي ـ أولاً ـ أن نرجع إلى جدول الأرقام العشوائية المُدرج في ملحق هذا الكتاب جدول رقم (1) ونبدأ في اختيارنا العشوائي من الصف السادس، العمود الخامس حيث نجد الرقم 42682، هذا الجدول يتكون من أعداد يتألف كل منها من 5 أرقام ولكننا، نحتاج إلى الرقمين الأوليين من كل عدد، ونبدأ في الاختيار، ونهمل الرقم الأول الذي يساوي 72، حيث أن قائمتنا لا تحتوي على أي أعداد أعلى من 50، وبقراءة الأعداد أسفل هذا العدد نقع على الأعداد 01921، ويظهر بعـد ذلك الرقم 21 مرة أخرى، ويتم إهماله أيضاً، حيث سبق له الظهور وتم اختياره بالفعـل ويُعد مكرراً، واختيارنا الأخير تكون الأرقام التالية في الترتيب 40,44، وعليه تكون عينتنا المُختارة هي الأسماء التي تحمل الأرقام: 48,40,21,12,1

وعـادة ما تكـون العينـات أكـبر حجـماً من ذلـك، ولكن أياً كان حجم العينـة، فالإجراءات واحدة لا تتغير. ويلاحظ أن كل المتطلبات المرغوبة في هذا الأسلوب عبارة عن قائمة قابلة للترقيم، وجدول أرقام عشوائي (يوجد الآن مراجع حاسوب تحتوي على إمكانات اختيار الأرقام عشوائياً، مما يعفي الباحث من القيام بهذه الإجراءات، ولكن عادة تستخـدم في حالـة الاختيار للعينات كبيرة الحجم). أما في حالة اختيار العينات يدوياً [2]، فيمكن استخدام أي عدد من الصفوف يراها الباحث ضرورية، وفي حالة عدم الحصـول على العـدد المطلوب من الحالات باستخدام أحد الصفوف يمكن مواصلة الإجراء باختيار صف آخر ونبدأ في اختيار الأعداد للوصول إلى عدد الحالات المطلوبة.

تُعـد جداول الأرقـام العشـوائية المعتمـدة (ونـظيراتهـا المـبرمجة بواسطة الحاسوب) إجـراءات مضمـونـة النتـائـج، طالمـا أننـا نتبع التعليمات المنصوص عليها في لائحة

(2) أي بدون استخدام الحاسوب (المترجمان).

الاستخدام، فالأرقـام موزعـة توزيعاً عشوائيا صحيحاً، ويمكننا الإعتماد عليها لاختيار عينات عشوائية علمية تصلح للمجالات البحثية المختلفة. وليس من المستبعد أن توجد بعض المشكـلات المتعلقـة بأعداد قوائم مجتمع الدراسة، ولذا عند تهيئة هذه القوائم لإجـراءات اختيار العينـة، يجب مراعـاة الآتي: ينبغي إدراج أي فرد من أفراد مجتمع الـدراسة (شخصاً كان أو كتاباً أو أي شيءآخر) مرة واحدة فقط على القائمة، وتكرار إدراج الفـرد أو الشيء مرة أخـرى على القائمة، يعني إعطاءة فرصة اختيار مضاعفة، ويُعد هذا تحيزاً مرفوضاً، لذا، فإننا نحتاج دائماً لمراجعة القائمة للتأكد من خلوها من هذه الأخطاء، وضمان تمثيلها الدقيق لمجتمع الدراسة. هناك مشكلات أخرى تتلخص في قِدم وعدم صلاحية القوائم، واحتوائها على أسماء لأشخاص متوفين، أو الغياب التام لأسماء أشخاص كان من الواجب إدراجهم، وعموماً، فإن الحصول على قائمة حديثـه ومنقحة وكاملة، أمر من الصعب تحقيقه، ونادراً ما نجد قائمة لا تحتاج إلى مراجعة لتصويب وتصحيح الأخطاء الناجمة عن الإهمال وعدم الصلاحية.

من منا لم يسمع بالإخفاق الذريع للتنبؤ بنتائج انتخابات ١٩٣٠ [3]حيث قررت مجلة شهيرة استدراج جمهور الناخبين واستمالة أصواتهم لصالح لاندون، واعتمدت في ذلك على قوائم دليـل الهاتف، وفـي هـذه الأيام، لـم يكن هناك كثير مـن النـاس يملكـون هواتف، وبالتالي لم يكونوا مدرجين في هذه القائمة، ويبدو أن أعداد الناخبين الذين لم يكن لديهم هواتف في ذلك الوقت، توافدوا وبكميات غفيرة على مراكز الاقتراع وانتخبوا فرانكلين روزفلت، وأدى هذا التنبؤ الخاطيء الذي قامت به المجلة لصالح منافسه، إلى توقفها عن الصدور. قـد يصـح أن الناخبين المالكين لأجهزة هواتف قد صوتوا لصالح السيد/ لانـدون، ولكن من الـواضـح أن أعداد بقية الناخبين غير المدرجين في قائمة الهاتف كانت أكبر منهم بكثير.

وحتى في وقتنا الحاضر، قد لانجد الطبقة محدودة الدخل مدرجة في قوائم الهواتف، وهـذه الـطبقـة قد تمثـل أهمية عظيمة، لبعض أنواع البحوث والدراسات. أما قوائم السلطات المحلية، فهي دائـماً قديمـة وغـير صالحة، حتى عند نشرها للمرة الأولى، وتحتاج - أيضا - إلى مراجعة وتنقيح، كذلك القوائم التي تحتوي على تكرار فإنها تصبح مشكلة حقيقة، إلا لو كان هـذا التكرار يمكن حصره وتصحيحه، ومثال لذلك، قوائم

(3) يشير المؤلف هنا إلى انتخابات الرئاسة في الولايات المتحدة الأمريكية عام ١٩٣٠، حيث كان المرشحان الرئيسيان فيها روزفلت و لاندون، وفاز فيها روزفلت بالأغلبية رغم أن التنبؤات التي أجريت قبل الانتخابات أكدت فوز لاندون (المترجمان).

العضوية، قوائم المنظمات، أدلة العاملين. والقوائم الشبيهة بها، فهي دائماً تحتوي على أسماء مكررة، ويستطيع أن يشهد بذلك كل من يتلقى مراسلات غير ذات معنى في بريده اليومي، نتيجة لتكرار اسمه على العديد من هذه القوائم. عموماً فإن هذه المشكلة، يمكن التغلب عليها بتنقيح وتصويب القائمة، بأفضل صورة ممكنة، وألا يتم استخدامها إلا إذا اطمأن الباحث تماماً على صلاحيتها للاستخدام.

إن استشارة الأشخاص الذين هم على علم بمجتمع الدراسة، يساعد إلى حد كبير في التغلب على كثير من هذه المشكلات، وفي بعض الأحيان يمكن معاودة النظر في مشكلات البحث، بحيث نستطيع موائمة مجتمع الدراسة مع القائمة المستخدمة. وتذكر دائما أن قائمة بحث متخمة بالأخطاء والتحيزات تكون أسوأ بكثير يمن عدم وجود قائمة، فغالباً ما تقود نوعيات القوائم السيئة سابقة الذكر إلى استنتاجات واستنباطات خاطئة.

المعاينة الإحصائية المنتظمة Systematic Sampling

تُعد المعاينة الإحصائية المنتظمة واحدة من أكثر أنواع العينات العشوائية شيوعاً وفائدة إذ أن محاولة إيجاد عينة عشوائية بسيطة من قائمة بالغة الطول لمجتمع دراسة أمر شاق للغاية (ولنا أن نتخيل محاولة إيجاد عينة من قائمة تحتوي على 150000 شخصاً، خاصة إذا كانت القائمة غير مرقمة)، ولكن باستخدام العينة المنتظمة، يمكن أن يتم هذا الأمر بسهولة.

وفي هذا النوع من العينات، بدلاً من استخدام جدول الأرقام العشوائية، فإننا نقوم بالاختيار العشوائي لنقطة البداية، ثم نواصل عملية الاختيار مع مراعاة ترك فواصل منتظمة بين الاختيارات المتتابعة (ويرمز للمسافة الفاصلة بين كل اختيار والذي يليه بالرمز K) مثال لذلك: إذا أردنا أخذ عينة تقدر بـ 40 بطاقة من فهرس بطاقات يتألف من 1000 بطاقة، ماعلينا في بهذه الحالة إلا اختيار البطاقات، وذلك بالاختيار العشوائي للبطاقة الأولى ثم ترك فواصل ثابتة بين الاختيارات تقدر بـ 25 [$25 = \frac{1000}{40}$] تتم هذه العملية كما يلي: نبدأ الاختيار باستخراج بطاقة بطريقة عشوائية من مجموعة الـ 25 بطاقة الأولى، ثم نبدأ في استخراج باقي البطاقات بعد ترك فواصل تقدر بـ 25 بطاقة لكل فاصل. ولنفترض أن البطاقة الأولى التي تم اختيارها عشوائياً كانت البطاقة رقم 10 وعليه سيكون باختيارنا التالي البطاقة رقم 35 وبعدها البطاقة رقم 60، والتالية لها رقم 85، . . . وهكذا حتى نصل إلى اختيارنا الأخير الذي سيكون البطاقة رقم 985، وبهذا نكون قد حصلنا على الـ 40 بطاقة المطلوبة بطريقة عشوائية منتظمة. من المهم

أن نبـدأ الاختيار الأول عشـوائياً، ويتم ذلك إما عن طريق استخدام جدول الأرقام العشوائي أو ببساطة أكثر نسأل زميل لنا أن يعطينا أي رقم ما بين 1 و 25 لنبدأ به. إن فرضية العشوائية، يمكن أن تتبع طالما لدينا قناعة بأنه لا يوجد في ترتيب القائمة شيء ما يجعلنا نفضل تطبيق الأسلوب غير العشوائي على المتغيرات المطلوب قياسها.

قد تنشأ مشكلة عند تطبيق أسلوب العينة العشوائية المنتظمة، فقد يحدث لسبب أو لآخر أو عن طريق الصدفة، أن تتطابق خصائص مفردات العينة المسحوبة عن طريق الفـواصـل المنتـظمة، حقاً، أن هذه المشكلة قد تكون نادرة الحدوث في مجال بحوث المكتبات، ولكن المثال التالي يوضح إمكانية حدوثها:

أراد أحـد المكتبيين، أن يُجرى مسحـاً ميدانياً يستـطلع فيه رأي مواطني المنطقة في الخدمات التي تقدمها مكتبته. وبغرض تحديد مجتمع الدراسة قرر استخدام قائمة يكون قوامها أرقام المنازل في المنطقة المحيطة بالمكتبة ليستطلع رأي قاطنيها، حيث أنهم يمثلون فئـة المستفيدين من المكتبـة، ثم بدأ بتـطبيق أسلوب العينـة العشـوائية المنتـظمة على القـائمـة، وتصـادف أن الـرقم العشوائي الأول الذي قام باختياره، وأرقام الفواصل المنتظمة التي ترتبت على هذا الاختيار، كانت ـ وبمحض الصدفة ـ تخص مواقع المنازل التي تقع على ناصية Corner Houses. ونتيجة لذلك تشكلت لدية «عينة من السكان» ذات وضـع اجتماعي واقتصـادي متميز عـن باقي سكـان المنطقة (4) وبالتالي فقد تختلف وجهات نظرهم تجاه الخدمات التي تقدمها المكتبة العامة، عن وجهات نظر باقي سكان المنطقة الذين يقطنون في منازل عادية وغير متميزة، مما سيؤثر ـ حتماً ـ على النتائج التي تنشأ من البحث.

عادة ما تستخدم القوائم المرتبة الفبائيا ـ مثل الببليوجرافيات ـ في كثير من العينات العشوائية المنتظمة في بحوث المكتبات، فهل هناك ـ حقاً ـ ما يثير الشك بأن في هذا الترتيب الألفبائي ما قد يخل بنظام العشوائية المتبع في تحديد المتغيرات المراد بحثها؟. وحقيقـة لا يخلو الترتيب الهجائي لأسماء المؤلفين أو المستفيدين من لمسة ذات طبيعة عرقية، حيث تتجمع أسماء العائلات المتشابهة ذات الأصول العرقية الواحدة في ترتيب

(4) يعني المؤلف أن المنازل تقع على ناصية ـ عادة ـ أسعارها مرتفعة، وبالتالي فقاطني هذه المنازل ـ ملاكـاً كانـوا أو مستـأجـرين ـ يتمتعـون بوضـع اقتصادي واجتماعي مختلف عن جيرانهم في المنطقة، وبالتالي تختلف ردود فعلهم تجاه الخدمات التي تقدم من قبل المكتبة العامة بالمنطقة. (المترجمان).

أبجدي متسلسل (مثال O's, Mc's) مما ينتج عنه عينة بها نسبة لا بأس بها من التجمع العــرقي، ولكن حتى إذا حدث ذلـك، فإنـه نادراً ما يسبب أي مشكـلات جدية لمستخدمي أسلوب العينات العشوائية المنتظمة في مجال المكتبات. وعلى أي حال، تُعد المعـاينة الإحصائية المنتظمة من أفضل الحلول، إذا أخذنا في الاعتبار القوائم بالغة الطول التي تستخدم في مجال بحوث المكتبات، فبافتراض أن علينا الحصول على عينة كبيرة الحجم، من قوائم تتضمن مئات الآلاف من بطاقات الفهرسة، فإن ترقيم كل بطاقة ومحاولة تشكيل عينة باتباع أسلوب جدول الأرقام العشوائية، يكون أمراً في منتهى الصعـوبـة، والحل الأمثل هو اتباع أسلوب العينة العشوائية المنتظمة، الذي سيوفر الكثـير من الـوقت والجهـد المبذولين. وهناك العديد من الطرق التي يطبق بها هذا الأسلوب (أسلوب العينات العشوائية المنتظمة)، فمثلاً يمكن قياس البطاقات بالبوصة أو بالسنتمتر (كم عدد البطاقات في البوصة أو في السنتيمتر)، ويتم تحديد الفواصل المنتظمة بمسافات تحدد بعدد البوصات أو السنتيمترات أو بأي مقياس طولي مناسب. والعيب الوحيد في استخدام أسلوب فواصل المسافات أن البطاقات القديمة، كنتيجة لكثرة الاستخدام تكون أطرافها بالية ومنحنية إلى أسفل، مما يؤثر على فرصة اختيارها حيث لا تكون ظاهرة للعيان.

وعلى أي حال، أياً كانت التقنية المستخدمة فمن الأهمية بمكان أن يقوم الباحث، بمراجعة وتحليل القائمة وتصويبها من الأخطاء قبل الاستخدام.

المعاينة الإحصائية الطبقية Stratified Sampling

في بعض الأحوال يكون في استخدام المعاينة الإحصائية الطبقية تطويراً لأسلوب العينة، إذ يعتمد أسلوب العينة الطبقية على تصنيف أو تقسيم مفردات مجتمع الدراسة إلى فئات متناسقة ومتفقة مع أغراض الدراسة. ومن البديهي أن كل مفرد من مفردات مجتمع الدراسة يظهر مرة واحدة فقط، وفي فئة واحدة فقط.

أما في أسلوب المعاينة الإحصائية الطبقية النسبية فيتم أخذ عينة عشوائية بسيطة أو منتظمة من كل فئة لنحصل في النهاية على مجموعة من العينات الفرعية ذات نسب متساوية بالنسبة لحجمها ودرجة تواجدها في مجتمع الدراسة.

فإذا قمنا ـ مثلاً ـ بعمل مسح لـ 100 مكتبة في إقليم ما، وكنا نعلم أن المكتبات المدرسية تمثل نسبة 50% ، والمكتبات العامة 30% ، والمكتبات الأكاديمية 15% ، و 5% مكتبـات متخصصـة، في هذه الحـالـة يكون لدينا المعلومات اللازمة لتصنيف هذه

المكتبات وتشكيل عينة طبقية، بحيث يحصل كل نوع من أنواع المكتبات على نسبة تساوي في مقدارها نسبة تواجده في مجتمع الدراسة (أنظر شكل رقم 9). (قد نرغب في بعض الأحيان تصنيف هذه الفئات بنسب غير متساوية في الحجم وسنوضح السبب وراء ذلك لاحقاً).

يمكن للتصنيف الطبقي المساهمة في تسهيل إجراءات المعاينة الإحصائية، فقد يتوافر لدينا عدد من القوائم المنفصلة، تحتوي كل منها على نوع محدد من المكتبات، عندئذ سيكون من السهل علينا أن نستخرج عينة من كل قائمة، بدلاً من تجميع القوائم كلها في قائمة واحدة ومحاولة تشكيل عينة عشوائية بسيطة أو منتظمة من القائمة الموحدة، واختيار العينة حسب الأسلوب الطبقي، يزيد من كفاءة العينة، بشرط التجانس الداخلي للفئات المكونة لمجتمع الدراسة، وبعبارة أخرى، إذا ما كانت المتغيرات في الخصائص موضع الدراسة متجانسة مع بعضها البعض، أكثر من تجانس الفئات المكونة لمجتمع الدراسة، فإن تطبيق أسلوب المعاينة الطبقية يكون هو الأفضل، أما إذا كانت الفئات المكونة لمجتمع البحث متجانسة بدرجة كبيرة (هذا يعني أن المتغيرات تكون أكثر حدة داخل كل فئة بينما تكون أقل حدة بين الفئات المختلفة) في

مجتمع البحث (100 مكتبة) / العينة (20 مكتبة)

العينة	مجتمع البحث	النسبة
10 مكتبات مدرسية (20 x ,50) مختارة عشوائيا	50 مكتبة مدرسية	% 50
6 مكتبات عامة (20 x,30) مختارة عشوائيا	30 مكتبة عامة	% 30
3 مكتبات أكاديمية (20 x ,15) مختارة عشوائيا	15 مكتبة أكاديمية	% 15
1 مكتبة أكاديمية واحدة (20 x , 5) مختارة عشوائيا	5 مكتبة متخصصة	% 5

شكل (9) رسم توضيحي للمعاينة الإحصائية الطبقية النسبية

هذه الحـالـة فإننـا لا نجني الكثـير، بل قد لا نجني شيئـاً، بتطبيق أسلوب المعاينة الإحصائية الطبقية.

إن الموقف يكون أكثر تعقيداً عندما نقوم بقياس متغيرات متعددة، فقد تكون بعض المتغيرات موزعة بالتساوي داخل مجتمع الدراسة (أي داخل كل الفئات المكونة لمجتمع الدراسة) بينما يتواجد البعض الآخر داخل بعض الفئات دون الأخرى، بمعنى أنها تتواجد داخل الفئات وليس بين الفئات، مثال لذلك:

إذا تم حساب الميزانيات بناء على تكلفة الفرد Per Capita، فإننا لا نجد فرقاً كبيراً في تلك التكلفة بين نوع أو آخر من المكتبات، وهذا يعني أننا لو حاولنا تطبيق أسلوب المعاينة الإحصائية الطبقية على هذه الحالة فلن نستفيد شيئاً. وعلى أي حال، فالشيء المهم حقاً، هو التأكد من وجود مفردات كافية (أي عدد كافي من المكتبات) داخل كل فئة (كل نوع من أنواع المكتبات) نستند عليه في تقييم المتغيرات التي تتواجد في الفئات المختلفة.

ويمكننـا في بعض الحالات ـ الحصول على النتائج التي نأمل الوصول إليها عن طريق تطبيق أسلوب المعاينة الطبقية غير النسبية. وفي المثال عاليه كان لدينا 5 مكتبات متخصصة فقط في قائمتنا، وأياً كان الأسلوب الذي سنتبعه سواء بتشكيل عينة عشوائية بسيطة أو منتظمة، أو عينة طبقية نسبية بواقع 30%، فلن نحصل في جميع الحالات إلا على مكتبة متخصصة واحدة فقط، وبالتالي لن تكون لدينا معلومات كافية عن هذا النوع من المكتبات، تعيننا على تحليل المتغيرات موضع البحث (مثل تكلفة الفرد في ميزانيات هذا النـوع من المكتبـات)، ولـذا فقـد نرغب في إضافة المكتبات الخمس المتخصصة ـ كعينة إضافية ـ لنضمن دقة وصحة النتائج التي نحصل عليها حول هذا النوع من المكتبات.

وقد يكون لدينا رغبة حقيقية في زيادة حجم عينة المكتبات المتخصصة أو العامة أو غيرها، إلا أن قرارنا هذا يتوقف ـ إلى حد كبير ـ على قناعتنا بأن المتغير المراد قياسه على علاقـة مبـاشرة بذلك النوع من المكتبات المطلوب زيادة عينته. فقد نرى مثلاً، أن ميزانيات المكتبات المدرسية متقاربة إلى حد بعيد (لا ضرورة في هذه الحالة لزيادة حجم عينة المكتبات المدرسية)، بينما نجد ميزانيات المكتبات المتخصصة ـ أو أي نوع آخر من أنواع المكتبات الموجودة في الدراسة ـ مختلفة ومتمايزة (وهنا قد نرى من الضروري زيادة حجم هذا النـوع من المكتبـات لنضمن وجود عدد كاف من المتغيرات، تقودنا بعد تحليلها إلى نتائج أكثر دقة).

إن الغرض من تطبيق أسلوب العينة الطبقية ـ كمثيله في العينة العشوائية البسيطة ـ هو محاولـة معـرفة مدى التغير الذي طرأ على المتغيرات التي نقوم بدراستها، وبعبارة أخـرى، ليس المقصـود بالمعـاينة الإحصائية العشوائية الطبقية الحصول على نسخة مصغـرة مطابقة لمجتمع الدراسة عن طريق التمثيل النسبي. [بخلاف ما يحدث في العينة الحصصية] [5] ويمكننا القول، بأن قرار استخدام أسلوب العينة الطبقية، يتوقف على ما إذا كانت الخصائص والسمات موضع البحث أكثر تجانساً داخل الفئات منها بين الفئـات ويتوقف كذلك على التكلفة التي تترتب على تطبيقه، وكما هو الحال مع كل الإجراءات المتعلقة بالمعاينة الإحصائية، فإن تكلفة المعاينة ترتبط ارتباطاً وثيقاً بمدى الدقة المطلوبة في النتائج، ولذا فإن قرارنا باتباع أسلوب معين للمعاينة الإحصائية، يتوقف على مدى استعدادنا لتحمل نفقات معينة للحصول على مستوى معين من الدقة في النتائج.

تصحيـح عـدم تناسب العينــات الطبقيــة:

Correcting Disproportionately Stratified Samples

تفيدنا العينة الطبقية غير المتناسبة، في تلخيص النتائج لكل فئة على حدة، بدون مشقة ما قد نعانيه في تصحيح العينة بالكامل. وعلى أي حال، فإن تلخيص العينة بأكملهـا، وبالأخص، تلخيص مجموعة الفئات المكونة للعينة، يتطلب منا، موازنة الإجـراءات بحيث نتمكن من تصحيح عدم تنـاسب التمثيل النسبي للفئـات التي تتكون منها العينة. وهذه الموازنة تصلح للتطبيق ليس فقط على العينات الطبقية. بل هي صالحة ـ أيضا ـ للتطبيق على العينات الحصصية (وهي إحدى أنواع العينات غير الإحتمالية).

وكمثال لذلك: دعنا نفترض قيامنا باستطلاع رأي 500 شخص بالغ من المجتمع المحيط بالمكتبـة، ما إذا كانوا يستخدمون المكتبة؟ وهل تخرجوا من المدارس الثانوية العليا؟. الجـدول التالي رقم (10) يلخص النتائج التي حصلنا عليها. (النتائج التي حصلنا عليها، تمثلها الأرقام غير الموضوعة داخل الأقواس).

دعنـا نفـترض مرة أُخرى، بأننا حصلنا على معلومات من الإحصاءات الرسمية للدولة، أو أي مصدر آخر موثوق به، تفيد بأن 60% من مجموع سكان هذه المنطقة

(5) العينة الحصصية إحدى أنواع العينات غير الاحتمالية، سيتم التعرض لشرحها لاحقاً (المترجمان).

جدول (10): بيانات خاصة بـ 500 شخص بالغ حول استخدام المكتبة، ومستوى تعليمهم

استخدام المكتبة	ليسو بخريجي مدارس ثانوية عليا	خريجي مدارس ثانوية عليــا	المجموع
نعم	10 (20)	200 (150)	210 (170)
لا	90	200	290
	100 (200)	400 (300)	500

حاصلين على الشهادة الثانوية العليا، وأن 40% من السكان غير حاصلين عليها، وهذا يخالف النسب التي حصلنا عليها من خلال عينتنا والتي تقدر بـ 80% للحاصلين على الشهادة الثانوية العليا، 20% لغير الحاصلين عليها وهذا يعني أن النتائج المتوقعة، والتي كان يجب أن نتوصل إليها، هي تلك الأرقام المدونة بين الأقواس داخل الجدول لو أننا حصلنا من خلال العينة التي قمنا بجمعها على النسب الحقيقة لأعداد الخريجين وأعداد غير الخريجين. بالإضافة إلى ذلك، فإن نتائج تحليلنا للعينة أشارت إلى أن نصف عدد خريجي المدارس الثانوية العليا (وعددهم 200) يستخدمون المكتبة (العدد الكلي للخريجين يقدر بـ 400) أما إذا استخدمنا الأرقام الحقيقة التي حصلنا عليها من الإحصاءات الرسمية، فسنجد أن الأرقام المتوقع الحصول عليها هي تلك المسجلة بين الأقواس مقابلة كلمة «نعم» أي: 20 بالنسبة لمن لم يتخرجوا من المدارس الثانوية العليا، و 150 لمن تخرجوا من المدارس الثانوية العليا (بفارق 50 عن نتائج العينة) ليصبح المجموع 170، وكنتيجة لذلك لو كنا أخذنا عينة متناسبة، وصححنا العينة التي حصلنا عليها لتطابق النسب المسجلة في الإحصاءات الرسمية، فيما يخص المستوى التعليمي لسكان المنطقة، لكنا وجدنا أن نسبة البالغين الذين يستخدمون المكتبة تساوي 34% بدلاً من 42%. وهذه النسبة الأخيرة المرتفعة (42%) حصلنا عليها بسبب أن العينة التي قمنا بجمعها لم تكن متناسبة تناسباً صحيحاً مع الأعداد الحقيقة للخريجين في مجتمع الدراسة الذي قمنا بتحليله.

كان لتصحيح الأرقام الخاصة بمستوى التعليم أثره الإيجابي على دقة التحليل والتقييم للمستفيدين من المكتبة من سكان هذه المنطقة، وحقيقة الأمر، أننا بإجرائنا هذا، قمنا بتصحيح التحيز أو التمثيل الخاطىء للعينة التي جمعناها، وبالتالي استطعنا أن نغير جذرياً النتائج التي حصلنا عليها سابقاً.

المعاينــة الإحصائيــة غير الاحتماليــة
NON-PROBABILITY SAMPLING

المعاينة الإحصائية الغرضية — Purposive Sampling

قد تسبب الاعتبارات الاقتصادية ـ في بعض الأحيان ـ عقبات تقف أمام تطبيقنا لأسلوب المعـاينة الإحصائية الاحتمالية، وترغمنا على اللجوء إلى استخدام عينة غير مرغوب فيها، تُعرف باسم العينة غير الاحتمالية، وربما تُعد العينة الغرضية من أكثر أنـواع العينات غير الاحتمالية أهمية، وقد يطلق عليها في بعض الأحيان اسم «عينة الخبرة». وتصلح هذه الطريقة في البحوث الاستكشافية التي غالباً ما نكون مهتمين فيها بالحصـول على معلومـات عن موضوع يكون معروفاً لدينا إلى حد ما، أولنا به خبرة سابقـة، ولضـمان حصـولنـا على معلومـات عالية الجودة، فعادة ما نلجأ إلى الخبراء المختصين بذلك الموضوع، أو إلى أشخاص لديهم خبرة واسعة وشاملة عن موضوع البحث، وذلك بدلاً من اللجوء إلى أسلوب الاختيار العشوائي من مجتمع دراسي كبير.

مثال بذلك: إذا ما أردنا إجراء بحث حول العلاقة بين المكتبات العامة والسلطات المحلية والتنفيذية في منطقة ما، فإن الحصول على أفضل معلومات حول هذا الموضوع، يأتي عن طريق استطلاع رأي مديري المكتبات المعروفين بنشاطهم، وفعاليتهم والذين يعون تماماً دورهـم الإجتماعي والسياسي في المنطقة، وذلك أفضل من استقاء المعلومات عن طريق عينـة من العاملين في المكتبات بالمنطقة. ومن الطبيعي، استحالة تعميم النتـائـج التي نحصـل عليهـا عن طريق هذا الأسلوب، خاصة فيما يتعلق بخبرات وقـدرات باقي مديري المكتبـات، ناهيك عن إمكـانية تعميمها على جميع المكتبيين بالمنـطقـة المعنية. أمـا إذا كانت لدينـا قنـاعتنـا الخـاصـة، بأن «عينة الخبراء» التي استخدمناها كافية لتمثيل مجتمع الدراسة، فبإمكاننا محاولة تطبيق نتائجنا على مجتمع الدراسة كله الخاص بمدراء المكتبات العامة المتميزين بقدرات قيادية فعالة. ويمكننا اختبار صلاحية تمثيل هذا النوع من العينات لمجتمع الدراسة المعني، باتباع أسلوبين.

الأول، نستخدم فيه النتائج التي حصلنا عليها بواسطة العينة الغرضية لإجراء استبيان أو مقابلات نطبق فيها أسلوب العينة العشوائية. والثاني يتلخص في قيامنا بالبحث عن أدلـة تؤيد النتـائج التي حصلنا عليها، عن طريق قراءة ومراجعة التقارير السنوية، وسجـلات المكتبـة، والانتـاج الفكري في مجال المكتبات، كما يمكننا اتباع أسلوب الملاحظة للتعـرف على سلوكيات مديـري المكتبات ومطابقتها مـع نتائجنا، فـإذا كانت

نتائج اختبارتنا مؤيدة لنتائجنا التي حصلنا عليها في السابق، فمن الممكن القول بأن العينة الغرضية التي استحدثناها تمثل مجتمع الدراسة الذي نقوم بدارسته.

المعاينة الإحصائية الحصصية Quota Sampling

تُعد المعاينة الإحصائية الحصصية، إحدى أنواع المعاينات الإحصائية غير الاحتمالية، وهي تهدف إلى إيجاد نسخة مصغرة من مجتمع الدراس تحتوي على بعض من خصائصه المميزة، مثال ذلك: إذا تبين أن 15% من المستفيدين من المكتبات هم طلاب التعليم الثانوي العالي، فإن الباحثين الميدانيين سوف يأخذون في الاعتبار هذه «الحصة أو النسبة» في اعتبارهم حينما يقومون بإجراء بحوثهم.

وبالرغم من أن المعاينة الإحصائية الحصصية تبدو في ظاهرها متشابهة إلى حد ما مع المعاينات الإحصائية العشوائية أو الاحتمالية - مع وجود بعض التحفظات على هذه المقولة - إلا أنها في جوهرها تختلف تماماً عن هذين النوعين من العينات. تستخدم المعاينة الإحصائية الحصصية بنجاح من قبل العاملين في مجال استفتاء الرأي العام، وأبحاث التسويق، وذلك لتحقيق أغراض متعددة. يجب ملاحظة أن مؤسسة [Gall up Poll](6) قررت إيقاف استخدام أسلوب العينة الحصصية، منذ عام 1972، وبدأت في تطبيق أسلوب المعاينة العشوائية).

المعاينة الإحصائية الصدفية Accidental Sampling

يمكن القول بأن المعاينة الإحصائية الصدفية متقاربة في مفهومها مع المعاينة الإحصائية الحصصية، فكلاهما يعني اختيار عينة بأسهل وأبسط الطرق الممكنة، مع عدم وجود ضمان تمثيل العينة المُختارة بهذا الأسلوب لمجتمع الدراسة المزمع دراسته، فعينة الصدفة للمستفيدين من مكتبة ما، قد تؤخذ عن طريق أول خمسين مستفيداً يتوجهون إلى موظف خدمة الإعارة. وعينة الصدفة لمواد مكتبة ما، قد تُشكل من أول خمسين كتاباً جديداً يرد إلى المكتبة عن طريق البريد، وقد تستخدم عينة الصدفة بطريقة أكثر دقة عندما يوضع لها بعض المحددات والمواصفات للخصائص التي يجب أن تتوافر في مفردات العينة.

مثال ذلك: يمكن الإرتقاء بمستوى عينة الصدفة، في حالة المستفيدين من خدمات المكتبة، فبدلاً من اختيار أول 50 شخصاً يتوجهون إلى مكتب الإعارة، نقوم باختيار

(6) مؤسسة استطلاع رأي تعمل على استفتاء الرأي العام، على مستوى الولايات المتحدة الأمريكية، والعالم في مختلف القضايا السياسية، الاجتماعية... الخ. (المترجمان).

أول 25 سيده، وأول 25 رجالاً يتوجهون إلى هذه الخدمة، فإذا إفترضنا أن عامل الجنس يحدد كيفية استخدام خدمات المكتبة، فإننا في هذه الحالة نستطيع الخروج بمؤشرات توضح لنا الاختلافات في استخدام خدمات المكتبة بين الذكور والأناث. أما في حالة المواد المكتبية، فنستطيع أن نضع بعض المتغيرات لتحسين نوعية العينة المطلوبة، بأن نحدد ـ مثلا ـ تاريخ نشر معين، أو مجال موضوعي معين (انسانيات/ علوم اجتماعية/ علوم بحتة . . . الخ)، ونستطيع في هذه الحالة اختيار أول 17 كتاب يرد إلينا بالبريد في هذه المجالات المحددة.

وعلى أي حال، لا يمكننا ـ باتباع هذا الأسلوب ـ أن نضمن صلاحية تمثيل العينة المختارة لمجتمع الدراسة المطلوب دراسته ـ فقبل كل شيء ـ وبصرف النظر عن جانب التكلفة والوقت، فإن الغرض الأساسي من المعاينة الإحصائية أن تمدنا بمجموعة من المفردات (العينات) تمثل مجتمع الدراسة تمثيلاً دقيقاً.

أن أفضل السبل للحصول على عينة تمثل مجتمع الدراسة تمثيلاً دقيقاً، يكون باتباع أسلوب المعاينة الإحصائية الاحتمالية، وإذا أردنا الدقة، فإن أفضل أنواع العينات التي يمكن أن نستخدمها ونطبقها، تتمثل في نوع خاص من أنواع المعاينة الإحصائية الاحتمالية، أشرنا إليه في السابق باسم العينة العشوائية الطبقية.

حجم العينة: **Sample Size**

تستخدم المعادلة التالية لتحديد حجم العينة

$$n = \left\{ \frac{z\,(\sigma_g)}{E} \right\}^2$$

حيث:

n = حجم العينة

σ_g = قيمة الانحراف المعياري (المخمَّنة)

E = الخطأ المسموح به

Z = درجة الدقة المطلوبة في وحدات الدرجة العيارية (Z)

ولتحديد حجم العينة الصحيح، ينبغي اتخاذ قرارات وتخمينات معينة. مثال ذلك: دعنا نفترض أننا بصدد تشكيل عينة تمثل أسعاراً لكتب يتم إستخراجها من عناوين كتب مسجلة في ملف طلبات المكتبة، وقررنا أن تكون دقة نتائجنا مساوية 95٪، وسمحنا لأنفسنا بنسبة خطأ قدرناها بـ 20 دولاراً، وذلك يعني، أننا على استعداد لقبول متوسط

حسابي للتكلفة يزيد أو يقل بمقدار 20 دولار عن المتوسط الحسابي الحقيقي للتكلفة الفعلية، ويبقي لنا تنبؤ آخر يتعلق بالإنحراف المعياري، ويمكن حساب هذا التنبؤ على ضوء خبراتنا السابقة، أو من خلال مراجعة تكلفة بعض الكتب الموجودة لدينا. وفي مثالنا هذا تم تقدير الإنحراف المعياري بمقدار 1,50 دولار.

وبتعويض المعادلة السابقة بالأرقام التي أوردناها في المثال، ينتج لدينا الآتي:

$$n = \left\{ \frac{(1,50)1,96}{20 \text{ دولار}} \right\}^2 = 216 \text{ حالة بحثية}$$

وتُعد نسبة الـ 95% التي وضعناها لحد الثقة (والتي تساوي 1,96 من وحدات الانحراف المعياري) نسبة مقبولة ويمكن اتخاذها كمعيار، أما إذا أردنا استخدام حدود ثقة أكثر دقة (99 % مثلاً)، فسيتطلب ذلك منا، زيادة أعداد الكتب المشكلة للعينة (سيصبح البسط في هذه الحالة 2,57 وحدة انحراف معياري بدلاً من 1,96) وسيتم شرح الثقة وحد الثقة بطريقة أكثر تفصيلا في الفصل التالي. ولقد حاولنا في هذا الفصل، التعريف ـ بطريقة سريعة ـ بتقنيات المعاينات الإحصائية ا لأكثر شيوعاً. ويمكننا القول ـ بوجه عام ـ أنه، إذا ماتم تحديد المتغيرات المرغوب فيها، واختيار حالات الدراسة، فإنه يمكننا بعد ذلك البدء في تجميع البيانات عن الحالات المطلوبة. فإذا كان عدد الحالات المجمعة يمثل المجموع الكلي لمفرادات مجتمع الدراسة، فإنه يمكننا استخدام أسلوب الإحصاء الوصفي لتحجيم وتأطير المعلومات المجمعة ووضعها في شكل ملخص إحصائي. أما إذا تم اختيار الحالات عن طريق تقنيات المعاينة الإحصائية الاحتمالية، فإن البيانات في هذه الحالة، يمكن استغلالها للخروج بمؤشرات تقديرية عن مجتمع الدراسة ككل، وذلك في حدود المتغيرات المرغوب في دراستها.

عندما نقوم بتعميم النتائج المتعلقة بقيم متغيرات مجتمع الدراسة التي توصلنا إليها من البيانات التي جُمعت عن طريق العينة الاحتمالية، نكون قد استخدمنا ما يسمى بالإحصاء الاستدلالي أو الاستنباطي والذي سنناقش بعضاً من نظرياته ومسائله المنطقية في الفصل الثالث.

المراجع والقراءات:

Blalock, Hubert M., Jr. *Social Statistics.* 2d ed. New York: McGraw-Hill, 1972, p. 213-215.

Bookstein, Abraham. "How to sample Badly," *Library Quarterly, 44, no. 2: 124-132.*

Mendenhall, William, Lyman Ott, and Richard l. scheaffer. Elementary Survey Sampling Belmont, Calif.: Duxbury, 1971.

Slonin, Mark. *Sampling.* New York: Simon and Schuster, 1960.

الفصل الثالث

الاحصـــاء الاستـنـبـاطي INDUCTIVE STATISTICS

الفصل الثالث
الاحصاء الاستنباطي
[1]INDUCTIVE STATISTICS

تقع التقنيات التي تعرضنا لها في الفصل الأول تحت تصنيف عام يُعرف بالإحصاء الوصفي، وهذه التقنيات تقوم بتلخيص مجموع البيانات بطرق متعددة، ولا تهدف في النهاية لأكثر من توصيف البيانات المعنية بالدراسة وذلك كما ورد في الفصل الأول.

يعطينا الإحصاء الاستنباطي فعالية أكبر في التحليل، إلا أن ذلك يكون على حساب تعرضنا لمخاطر معينة. ويستند الإحصاء الاستنباطي ـ الذي يعرف أيضاً باسم الاحصاء الاستهلالي أو الاستنتاجى (Inferential) ـ على الأسباب التقليدية والمنطقية للتنبؤ، بمعنى قيامنا بتعميم المعرفة التي نحصل عليها عن طريق تحليل جزئية محدودة من البيانات على كلية البيانات موضع الدراسة، وباختصار، فالتنبؤ يعني استنباط الكل من الجزء.

ولأسباب كثيرة (عادة ماتكون قلة الإمكانات المادية، أو أن البدائل الأخرى غير متاحة)، قد نكون مرغمين على بحث مجتمع الدراسة موضع اهتمامنا عن طريق عينة مسحوبة من هذا المجتمع، وفي هذه الحالة، فالخصائص الكمية للعينات يطلق عليها الإحصاءات، وتسحب العينات من مجتمع الدراسة بغرض تقدير الخصائص المميزة والمطابقة لهذا المجتمع، وتختلف قيمة الإحصاءات من عينة إلى أخرى، بسبب إمكانية وجود اختلافات في أسلوب جمع العينات من مجتمعات الدراسة المتباينة.

تعرف الخصائص الكمية لمجتمعات الدراسة باسم معلمة Parameter[2]، وبسبب إستخدامه في جميع الحالات البحثية في عمليات التحليل الرياضي، فإن قيمة المعلمة تكون ثابتة لا تتغير.

---

(1) يطلق عليه ـ أيضاً ـ في بعض المراجع الإحصائية: الإحصاء الاستدلالي أو الإحصاء التحليلي. (المترجمان).

(2) تعني الثابت الذي لا يغير قيمته حسب المجال الذي ينتمي إليه (المترجمان).

خلال مناقشتنا للعينات ومجتمعات الدراسة، نستخدم مصطلحات متشابهة إلى حد كبير من حيث الشكل والصياغة، مثل ذلك: المتوسط الحسابي (الوسط الحسابي)، النسب. . . وغيرها. وعلى أي حال، عند استخدام هذه المصطلحات لتوصيف العينات، فإن مغزاها ومفهومها يختلف عما إذا استخدمت لتوصيف مجتمعات الدراسة، وأهم المصطلحات التي تعنينا في هذا الجزء من الموضوع، المتوسط الحسابي، الذي يرمز إليه بالرمز $\bar{X}$ عندما يكون على علاقة بإحصاءات العينة، كما يرمز إليه ـ أيضاً ـ بالرمز (μ)، كمعلمة للبحث (أو كمتوسط حسابي لمجتمع البحث)، والمصطلح الآخر هو: الانحراف المعياري، الذي يرمز إليه *u (Mu)* لرمز "S" عندما يرد في إحصاءات العينة في حين يرمز إليه بالرمز (Sigma)، عندما يرد كمعلمه للبحث (كانحراف معياري لمجتمع الدراسة بأكمله).

الفروض الإحصائية The statistical Hypothesis

تُعد الفروض ـ بوجه عام ـ تفسيراً لحدث ما أو عدة أحداث مجهولة، وعندما يقوم الباحث باختبار الفروض، فإنه يعمل على تشكيل فروض تتضمن تخمينات معينة حول القيم أو العلاقات المتبادلة بين مختلف المعلمات [3] الخاصة بمجتمع الدراسة. يقوم بعدها الباحث بالتأكد من نتائج عينته ومدى مطابقتها أو مخالفتها للتخمينات التي وضعها، ولذا فإن الفروض تعد ـ أساساً ـ شكلا من أشكال التخمين وينقسم الإحصاء الاستنباطي إلى نوعين من الفروض، أولهما يسمى بفروض البحث وثانيهما يدعى الفروض الإحصائية أو الفروض الصفرية Null Hypothesis [4] تتشكل فروض البحث من مزيج من التخمين والشعور الحدسي حول شيء ما، وتصاغ في شكل جملة تقريرية، مثال ذلك: يمكننا افتراض أن المستوى التعليمي للمستفيدين، يرتبط ارتباطاً مباشراً بكيفية استخدامهم للمكتبة، وهذا يشكل فرضية بحث، تستند على شعور حدسي قوي يشير إلى أن الأشخاص الأوفر تعليماً، هم الأكثر قراءة للكتب، وقد يستند شعورنا الحدسي، على أمر نعتقد ـ حقاً ـ في صدقه، وكنتيجة لذلك، نعمل على إثبات فروضنا، لكن في البداية، دعنا نؤكد مبدأ رئيسي ومهم، وبدون الدخول في المسائل اللغوية والتجريدية والمعضلات الرياضية المتعلقة بإثبات الفروض، طالما أننا استخدمنا عبارة « إختبار الفروض Test of Hypothesis » بدلاً من عبارة «إثبات

---

(3) جمع معلمة.

(4) يطلق عليها أيضاً فروض العدم.

الفروض Proof »، فإننا بذلك نكون قد وضعنا الفروض الإحصائية في موضع أقرب لنفيها منه لإثباتها، وتذكر، أن اختبار الفروض لا يعني إثبات أي شيء، فهو ـ بالكاد ـ يستطيع تزويدنـا بأدلة لتعضيد أو نفـي النظرية التـي ننادي بها. وتتطلب منـا الأمانة والدقة العلمية، أن نبذل قصارى جهدنا للتأكد من اتباعنا الموضوعية وعدم التحيز في الإجراءات المتعلقة بجمع معلومات البحث وقياسها وتحليل بياناتها، وبصفة بخاصة عند قيامنا بتشكيل مفاهيم ومشكلات البحث وفروضه. والفروض الإحصائية أو كما تسمى أحياناً الفروض الصفرية، هي الوسيلة الفعالة والمباشرة والسهلة أيضاً، التي تساعدنا على إجراء البحث بأمانة علمية.

ولذلك فإن الفروض الصفرية للمثال الذي أوردناه قبل قليل «في الفصل السابق»، ستنص على عدم وجـود علاقة بين مستوى التعليم للمستفيدين وكيفية استخدامهم للمكتبـة، وبالطبع، فإننا بإدراكنا الحسي، أو تخميننا أو قناعتنا (المتمثلة في فروض البحث)، يمكن أن نأمـل بإثبـات عدم صحة الفروض الصفرية. إن المنطق وراء استخـدام هذا الأسلوب، هو أن نفـترض عدم وجـود علاقـة بين مستـوى التعليم واستخدام المكتبة، ثم إذا ما ثبت ـ بالأدلة الدامغة ـ المستمدة من البيانات التي قمنا بجمعهـا بأن هنـاك علاقـة فعلية، فإننـا نستطيع استنتاج أن فروضنا الصفرية غير صحيحة، وعليه، فسنجد ما يدعم نظريتنا بوجود علاقة بين متغيرات مجتمع الدراسة موضـع اهتـمامنا، بالاستناد على عينة البيانات التي قمنا بجمعها. بالإضافة إلى أن اتباعنا لأسلوب الفروض الصفرية، سواء بإثباتها أو بنفيها، يجب أن نطرح على أنفسنا، تساؤل محدد: ماهي الحقائق الأخرى المتعلقة بالبحث؟، إن الفرض الجيد، هو الذي يقودنا بطريقة عملية للتوصل لفروض أخرى قد تكون مخالفة لفروضنا التي وضعناها سابقاً.

من خلال سعينـا للبحث عن الحقيقـة، والمسببـات، أو عن التنبؤات للعلاقات المتبادلة أو المكونة للمتغيرات، يجب أن نتبـع أسلـوب التساؤل المستمر، لنعيد صيـاغة المفـاهيم، والقنـاعـات، ونحـاول التفكـير في إيجاد بدائل جديدة للفروض المتعلقة بمشكـلات ومعـطيات البحث الـذي نقوم به، ومن الضروري للغاية أن نضع في اعتبارنا، المتغيرات المتعددة التي يمكن أن تطرأ وتؤثر على الفروض الموضوعة للبحث.

وعند استخدامنا للإستنتاج الإحصائي [5]، يجب أن نحـذر ثـلاث مسببات رئيسية

(5) هنا استخدم المؤلف مصطلح Statistical inference

(قد تكون متعلقة بمصادر البيانات أو عينات وفئات البحث) للأخطاء. الأول: مهما كانت الاحتياطـات المتخـذة حيال الإجراءات والاستنتاجات الرياضية، فمن المهم جداً، التـأكد من أن المعايير المتبعة موثوق بها وصالحة للتطبيق قدر الإمكان. فمهما كانت المعالجات الرياضية دقيقة وبارعـة فهى تكون عديمـة القيمة، إذا ما كانت معايير البحث غير صحيحـة، وكنتيجـة لذلـك، فقبـل الشروع في تحليل البيانات، يجب الحرص على مراجعة أساليب جمعها ومعالجتها بدقة شديدة، وذلك بغرض التأكد من خلوهـا من أي أخـطاء، والثاني: يجب التأكد من صلاحية طبيعة العينات، وطرق تشكيلها، وتناسبها مع الأسلوب الاستنباطي المستخدم. وفي كل الحالات المذكورة في هذا الكتاب، سنجد أن المعاينة الإحصائية العشوائية المستقلة Independant Random Sampling، صالحـة للتـطبيق مع اختبـارات الفـروض والإجـراءات الاستنبـاطية الأخرى، وقد تم مناقشة خصائص المعاينة الإحصائية العشوائية في الفصل السابق المتعلق بالمعـاينــة الإحصـائية، والثـالـث: من الضروري اتبـاع كل المتـطلبـات، والفروض، والإجراءات الأخرى وتطبيقها على كل الأساليب المتبعة في البحث، ومثال ذلـك: تتطلب بعض نماذج اختبارات الفروض أن يأخذ توزيع المتغيرات في مجتمع الدراسة شكلاً خاصاً. فاختبار (t) ـ مثلاً ـ يتطلب أن يكون توزيع المتغيرات في مجتمع الدراسة توزيعاً طبيعياً. إضافة إلى ذلك يجب أن يجري اختبار (t) على بيانات مقاسة على المستـوى الفاصل لمنحنى التوزيع التكراري، لنتمكن من الحصول على بيانات قابلة للتفسير، ويُظهر جدول الإجراءات الإحصائية (جدول رقم (1))، الاختبارات وما يناسب كل منها من مختلف مستويات القياس.

لن نتعرض في هذا الكتاب، إلى المبادىء الرياضية لتناسب مستويات القياس، ولا للنماذج الخاصة باختبار الفروض، ولكن من يرغب في الاستزادة حول هذا المواضيع، عليه بمراجعة الإنتاج الفكري في مجال الإحصاء الرياضي والتطبيقي.

والحـدس، أو البـديهـة، هو الشيء الذي يمكن أن يقود الشخص للتعرف على الاختيار الصحيح واختباره، أما الخبرة التي تكتسب عن طريق الاطلاع على الإنتاج الفكري الذي يعالج هذه الموضوعات، فهي تساعد ـ فقط ـ في تعزيز وفهم المتطلبات الصحيحة لاختيار المعايير وطرق الاختبار. إن استخدام المعايير والاختبارات المناسبة، أمر في غاية الأهمية، وإلا فإن كل الجهد الذي يبذل في تجميع البيانات سيضيع هباء، نتيجة للتحليل الخاطيء.

ويستـوجب اختبـار الفـروض استخـدام ثلاثة أنواع من التوزيعات المرتبطة مع

بعضها: توزيع العينة (بيانات العينة)، وتوزيع المعاينة الإحصائية (التوزيع النظري لإحصاءات العينة)، وتوزيع مجتمع الدراسة (توزيع جميع الحالات في مجتمع الدراسة على المتغيرات موضع الاهتمام). وتعد توزيعات المعاينة الإحصائية، وتوزيعات مجتمع الدراسة، توزيعات نظرية، وبالرغم من أن كلاهما يمكن تشكيله، إلا أننا من النادر مانهتم بالقيام بذلك. وعلى أي حال، أن استيعابنا لمفهوم توزيع المعاينة الإحصائية، ضروري للغاية، حتى نتمكن من فهم الأسس المنطقية لإجراءات اختبارات الفروض.

لنتخيل أننا بصدد تشكيل عينة واحدة، مكونة من خمسة مستفيدين، لمحاولة تحديد المتـوسط الحسابي لعدد الكتب التي استعاروها خلال السنة الماضية. وتتكون حالتنا البحثية هذه، من عدد من الأفراد المستخدمين للمكتبة، والمتغير المطلوب قياسه، هو عدد من الكتب. وغالباً، ما يمثل المستفيدون من المكتبة مجتمع دراسة كبير، يسمح لنا بتشكيل العديد من العينات البحثية، فلو قمنا بتكوين عدة عينات فإننا سنحصل على عدة متوسطات حسابية للكتب المستعارة، ولذا فسنجد لدينا: $1\overline{X}, 2\overline{X}, 3\overline{X}$ . . . الخ. أي عدد من المتوسطات الحسابية مساوياً لعدد العينات التي قمنا بتشكيلها، ويُلجأ في هذه الحالة إلى توزيع هذه المتوسطات الحسابية على رسم بياني، لنحصل على تمثيل بياني لتوزيع المعاينة الإحصائية للمتوسطات الحسابية. وإذا قمنا بسؤال كل المستفيدين من المكتبة، عن عدد الكتب التي قاموا باستعارتها من المكتبة خلال السنة الماضية ثم قمنا بجدولة البيانات التي حصلنا عليها، نكون بهذا قد قمنا بعمل توزيع للمتغير، ـ الذي يُعد في هذه الحالة ـ توزيعاً لمجتمع الدراسة. وبنفس المفهوم، إذا استخدمنا البيانات التي حصلنا عليها من عينتنا الأصلية المكونة من خمسة مستفيدين فقط، ثم قمنا بعمل توزيع تكراري لهذه البيانات، نحصل ـ بهذه الطريقة ـ على مايسمى بتوزيع العينة. ومن الطبيعي، أن اختبارات الفروض، لا تقتصر فقط على المتوسطات الحسابية أو على توزيع المعاينة الإحصائية للمتوسطات الحسابية، بل يمكننا تشكيل واختبار الفروض الإحصـائية ـ تقـريبا ـ حول أي نوع من أنواع الإحصاءات تكون موضع اهتمامنا. فمثلاً: إذا ما كانت فروضنا الإحصائية تتعلق بالانحراف المعياري لمجتمع الدراسة سنحتاج هنا لفحص توزيع المعاينة الإحصائية للانحراف المعياري للعينة، فإذا كان اهتـمامنـا منصباً على العلاقات التبادلية لمجتمع الدراسة، فسوف نحصر اهتمامنا في توزيع المعاينة الإحصائية لمعامل الارتباط للعينة.

اختبار الفروض للمتوسط الحسابي لمجتمع الدراسة

Hypothesis Testing about the Population Mean

دعنا نلقي نظرة فاحصة على أسلوب اختبارات الفروض، ونلخص بعض الأفكار الرئيسية التي وردت حول هذا الموضوع:

أولا وقبل كل شيء، بعد قيامنا بتحديد فروض البحث، نعمل على وضع خطة للبحث حتى نضمن أن نتائجنا لن تتأثر بأسلوب المعاينة الإحصائية الذي سنقوم بتطبيقه، وكنتيجة لذلك، نضع خطة لمعاينة أحصائية عشوائية بغرض الحصول على عينة تمثل مجتمع الدراسة تمثيلاً دقيقاً، بعدها نقوم بوضع فروض صفرية لمعلمات مجتمع الدراسة، ثم نشرع في تجميع وتحليل بيانات العينة. وحتى يسهل متابعة المناقشة القادمة، دعنا نفترض أننا نهتم بتقييم المتوسط الحسابي لمجتمع دراسة، حيث تنص الفرضية الصفرية على أن قيمة المتغير للمتوسط الحساب لمجتمع الدراسة تساوي 30 ، ودعنا نفترض أن الفرضية الصفرية صحيحة وهذا يعني أن المتوسط الحنسابي لمجتمع الدراسة يساوي 30 بالفعل. ولكن، حتى لو كانت الفرضية الصفرية صحيحة، فإن العينة الإحصائية (عينة المتوسطات الحسابية) التي سيتم حسابها، على ضوء العينات العشوائية المسحوبة من مجتمع الدراسة سوف تتغير قيمتها من عينة إلى أخرى، بناء على قانون الصدفة أو على الانحراف العشوائي، بسبب أن أي عينة سنقوم باختيارها سوف تتشكل ـ حتماً من عدد محدود من الحالات، يقل كثيراً عن العدد الفعلي للحالات الكلية لمجتمع الدراسة.

ولذا، ستتوقف قيمة المتوسط الحسابي للعينة، على عدد الحالات التي ضُمنت في العينة المختارة، فإذا أخذنا في الاعتبار كل القيم الممكنة للمتوسط الحسابي للعينة، فسنجد أن احتمالية ظهور أو تكرار بعض هذه القيم يكون كبيراً.

وفي حالة المتوسط الحسابي للعينة، فإن نسبة كبيرة منها ستكون متقاربة للغاية مع القيمة الحقيقية للمتوسط الحسابي لمجتمع الدراسة، والتي تساوي 30. على أي حال سيتكـرر ظهور بعض هذه القيم بصفة مستمرة، أما البعض الآخر، فسيكون نادر الظهور، وعليه، فقد نحصل على عينات قليلة للغاية، تكون متوسطاتها 4 أو 50 ، وسيظهر لنا التوزيع التكراري لكل هذه المتوسطات الحسابية المحتملة للعينة، أياً من القيم أكثر احتمالاً للظهور (أي متقارب نسبياً مع المتوسط الحسابي لمجتمع الدراسة)، وأياً منهـا ستكـون نادراً (أي بعيداً كل البعـد عن القيم المفترضة للمتوسط الحسابي

لمجتمع الدراسة) وعادة، عندما نقوم بتعريف القيم نادرة الظهور، فإننا نختار القيم التي تظهـر ما بين 1% ,5% ، وتبعـد عن القيم المفترضة للمتوسط الحسابي لمجتمع الدراسة، ويُعد التوزيع الكلي للقيم والذي يسمى توزيع المعاينة الإحصائية توزيعاً نظرياً، لأنه لا يتم في التطبيق العملي - مطلقا - بأن نأخذ عدداً غير محدد من العينات من مجتمـع الدراسة [هناك دائماً حدود في حجم وعدد العينات المختارة]. وبذلك، فنحن نأخـذ قيم المتـوسط الحسـابي لعينتنا، عن طريق حسابها من العينة التي قمنا باختيارها عشوائياً من مجتمع الدراسة، ثم نقوم بمقارنتها مع التوزيع النظري لكل المتوسطات الحسابية للعينة، والتي كان من الممكن حسابها من عينات بنفس الحجم، ومـأخـوذة من نفس مجتمع الدراسة، فإذا نتج لدينا متوسط حسابي للعينة، يصنف كمتوسط حسابي نادر الظهور، فهو يجعلنا نستنتج أن هذا المتوسط الحسابي غير محتمل الظهور في مجتمع بحثي متوسطه الحسابي يساوي 30 ، وهنا نستبعد الفرضية الصفرية التي وضعناها، خلاصة القول، نحن على إستعداد لتقبل الفرضية الصفرية للبحث، كحقيقة واقعة، إلى اللحظة التي يتضح لنا فيها، بالدليل القاطع، وعن طريق العينة التي قمنا بتحليلها، أن هذه الفرضية غير صحيحة. ولهذا فنحن نستند في اتخاذنا أي قرار يخص معلمة مجتمع البحث على عينتنا الإحصائية.

في اتخاذنا لقرار رفض الفرضية الصفرية للبحث، فإننا نتحمل مخاطرة أن يكون قرارنا هذا خاطئاً، وعادة ما تأخذ قيمة المخاطرة التي نود تحملها شكل النسبة المئوية، ومستويات النسبة المئوية التقليدية في هذه الحالة تكون مابين 1% و 5% ، ولكن لا يجب أخذ هذه النسب على أنها ثوابت.

ترتبط هذه النسب أو المخـاطرة التي نحن على استعداد لتحملها، بتعريفنا لقيم العينة نادرة الظهور، والتي تقدر قيمتها بأقل من الفرضية الصفرية، وينبغي أن نتذكر أنه بالرجوع إلى مثالنا السابق؛ نجد أن المتوسطات الحسابية للعينة التي قدرت بـ 4 أو 50 تُعد من المتوسطات أو النسب نادرة الحدوث، عندما يكون المتوسط الحسابي لمجتمع الـدراسـة مسـاوياً 30. لذا فعنـدمـا يكون المتوسط الحسابي لعينتنا بعيداً عن القيمة الافتراضية للمتوسط الحسابي لمجتمع الدراسة التي تساوي 30 ، وعندما تظهر لنا قيم المتوسطات بنسبة 5% فقط، فإن نتائجنا تكون مؤشرا على أن المتوسط الحسابي لمجتمعنا الدراسي لا يساوي 30. ولكن، قرارنا هذا قد يكون - أيضاً - خاطئاً. وللتأكد من ذلك يجب إعادة اختبار فروضنا أكثر من مرة فإذا ثبت بطريقة قاطعة أن المتوسط الحسابي لمجتمعنا الدراسي يساوي 30 فعلاً، فإننا نستبعد فرضنا البحثي المتعلق بالخمسة في

المائة أو قد يحدث العكس. وعموماً، فإن الهدف الذي نسعى إليه، هو الإقلال بقدر الإمكان من احتمال اتخاذنا للقرار الخاطىء أو الذي يترتب عليه رفضنا لفرض قد يكون صحيحاً.

لنفترض أن العينـة التي سحبنـاها من مجتمع الدراسة، تتكون من 35 مستفيداً والجدول رقم (11) يحتوي على توزيع العينة (أو توزيع القيم التي تحتوي عليها العينة) لهؤلاء المستفيدين، وعلى عدد الكتب المستعارة خلال السنة الماضية.

عدد الكتب	التكرار
100	2
10	10
5	15
4	5
0	3

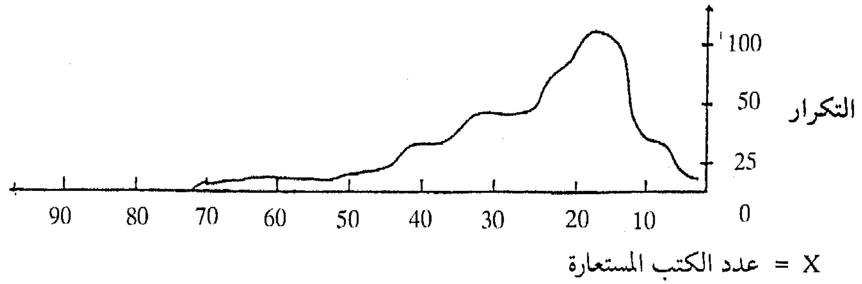
$$(\bar{X}) \text{ المتوسط الحسابي للعينة } = \frac{[(3 \times 0) + (5 \times 4) + (15 \times 5) + (10 \times 10) + (2 \times 100)]}{35}$$

$$= \frac{395}{35}$$

$$= 11{,}3$$

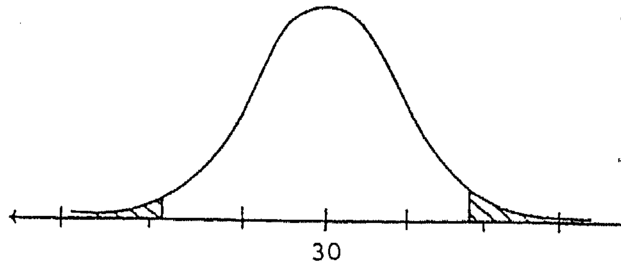
تقدر قيمة المتوسط الحسابي للكتب المستعارة، لهذه المجموعة بـ 11.3 ، والشكل التوضيحي رقم 10 يوضح التوزيع التكراري لمجتمع المستفيدين بالكامل، على ضوء عدد الكتب المستعارة، مع ملاحظة أن المتوسط الحسابي الحقيقي لمجتمع الدراسة لم يتم حسـابه. فإذا افترضنا أن المتوسط الحسابي لعدد الكتب المستعارة للسنة الماضية بالقياس مع مجتمع الدراسة كاملاً يقدر بـ 30 فإن القيمة النظرية لتوزيع المعاينة (توزيع كل المتوسطات الحسابية الممكنة للعينة، عندما n = 35) يمكن أن تأخذ الشكل الذي يظهر في الشكل البياني رقم 11 ، وقد يحدث أن تحصل بعض عينات الـ 35 مستفيداً على متوسط حسابي منخفض يقدر بـ 5 كتب مستعارة، والبعض الآخر قد يحصل على متوسط حسابي مرتفع قد يصل إلى 40.

وبالنظر إلى شكل (11) ، نجد أن القيم نادرة الظهور [أو المتوسطات الحسابية الناتجة عن عينـات، وتكـون قيمتهـا بعيدة كل البعـد عن قيمـة المتوسط الحسابي لمجتمع



شكل (10): التوزيع التكراري النظري للمجتمع الكلي للمستفيدين من المكتبة، على ضوء العدد المتغير للكتب المستعــارة

الدراسة]، توجد في أطراف توزيع المعاينة، وإن القيم المسجلة على الجانب الأيمن أو الجانب الأيسر، أقل ظهوراً من القيم المسجلة في القسم الأوسط من التوزيع، ولذا فإن ارتفاع التوزيع في الأطراف أقل كثيراً من ارتفاعه في الوسط. ويمكننا تحديد القيم في الأجزاء الطرفية من شكل (11)، التي تُعد من القيم النادرة الظهور، وذلك بوضع حدود فاصلة بينها وبين الأجزاء الأخرى، ثم تظليلها بخطوط متوازية. وهذه المناطق المظللة تمثل نسبة مئوية معينة من المساحة الكلية التي يمثلها المنحنى البياني (وبعبارة أخرى، نسبة معينة من العدد الكلي للعينات المسحوبة من المجتمع الدراسي)، ويمكن وضــع هذه النسب في صورة مئـوية (مثــال 5% , 2.5% لكل طرف)، وهذه المنطقة (المظللة) تمثل مساحة المخاطرة والنسب المئوية تمثل الوقت (على المدى الطويل) الذي سنرفض فيه الفرضية الصحيحة.



$\bar{X}$ = قيمة المتوسط الحسابي للعينة

القيمة الافتراضية للمتوسط الحسابي لمجتمع البحث = 30

شكل (11): توزيع المعاينة الإحصائية النظري، للمتوسط الحسابي لعدد الكتب المستعارة، من قبل المستفيدين، مستنداً على عينات عشوائية مكونة من 35 مستفيداً

وينبغي أن نتـذكـر أننـا افـترضنا حصولنا على المتوسط الحسابي الحقيقي لمجتمع الـدراسـة، وأن توزيع المعاينة الإحصائية تحدد عن طريق الفرضية الصفرية (الذي سيناقش عندما نتعرض لشرح نظرية النهاية المركزية The Central-Limit Theorem) ولذلك، فقد نحصل ـ في بعض الأحيان ـ على قيم مختلفة للعينات، أو على الأقل بعيدة كل البعـد عن المتـوسط الحسابي الحقيقي لمجتمع الدراسة حتى عندما تكون الفرضية الصفرية صحيحة. ولكون هذه القيم بعيدة عن هذا المتوسط، فإننا ـ عادة ـ ما نتخذ قراراً خاطئاً، برفض الفرضية الصفرية.

وحتى هذه اللحظة، كان تحديدنا غير واضح فيما يتعلق بالقيم التي نعتبرها قليلة أو نادرة الظهور، ويمكننا أن نقوم بهذا التحديد، بأن نشرع أولاً باختيار حجم عينتنا من خلال المعلومات التي نحصل عليها من نظرية النهاية المركزية، يمكننا استخدام قيمة درجة ا لانحراف المعياري Z Score، لتحديد القيم المتقطعة للمناطق المظللة. ودعنا الآن، نقوم بشرح نظرية النهاية المركزية، ثم نستعرض إجراءاتها بالكامل خطوة بخطوة وذلك بالإستعانة بالأمثلة.

نظرية النهاية المركزية The Central-Limit Theorem

يمكننا ملاحظة أن توزيع المعاينة الإحصائية للمتوسط الحسابي في شكل (11)، يتمركز حول المتوسط الحسابي الافتراضي لمجتمع الدراسة، الذي يقدر بـ 30، ويبدو كتوزيع طبيعي، وذلك بفضل المعلومات التي حصلنا عليها عن طريق نظرية النهاية المـركـزية. تزودنـا هذه الـنظرية الـرائعـة، بالقـاعدة الشرعية، لإستخدام التوزيع الطبيعي، كنموذج استنباطي:

1 – عندما نقوم باختبار الفروض المتعلقة بالمتوسط الحسابي لمجتمع الدراسة.

2 – عندما يكون حجم العينة كبيراً نسبياً.

ونحن قادرون ـ تماماً ـ على اختبار فروض العلاقات المتبادلة بين معلمات مجتمع الدراسة، أو للمعلمات الأخرى للمجتمع نفسه، ولكن النماذج الرياضية والتوزيعات النظرية للمعاينات الإحصائية التي نستخدمها في اختبار الفروض، قد لا تتطابق مع التـوزيع الـطبيعي، وهنـا نحن نهتم فقط بالفـروض المتعلقـة بالمتـوسط الحسـابي لمجتمع الدراسة.

ولو قمنا بتشكيل عدد من العينات العشوائية المستقلة ذات الأحجام المحددة من مجتمع دراسة ما، وكان المتوسط الحسابي يساوي (μ)، التغاير يساوي (σ^2)، فإن

ذلك يمكننا من تشكيل توزيع معاينة إحصائية من المتوسطات الحسابية للعينة، الذي يُحسب على ضوء العينات العشوائية المستقلة.

وتنص نظرية النهاية المركزية، على ما إذا كان حجم العينة كبيراً نسبياً (أكثر من 30 مثـلاً)، فإن توزيع المعاينة الإحصائية للمتوسط الحسابي) ($\overline{X}$) يصبح ـ تقريباً ـ موزعاً توزيعاً طبيعياً، ومع متوسـط حسابي (لمجتمع الدراسة) مقـداره (μ)، وتغايـر مقداره ($\frac{\sigma^2}{n}$)، يقـترب توزيع المعـاينة الإحصائية من التوزيع الطبيعي. ويلاحـظ، أن توزيع المعـاينـة الإحصائية العشوائية للمتوسط الحسابي للعينات كبيرة الحجم نسبياً، سيكـون ـ تقـريبـاً ـ موزعـاً توزيعاً طبيعياً، بصرف النظر عن شكل التوزيع لمجتمع الدراسة على ضوء المتغير موضع اهتمامنا.

وبصرف النـظر عن الشكـل، يمكننـا افـتراض أن التـوزيع للمعـاينة الإحصائية العشـوائية سيأخذ ـ تقريباً ـ الشكل الطبيعي، ولذا، يمكننا استخدام المعلومات عن مجموعة التوزيعات الطبيعة، لتحديد ماهية النسب المئوية للمساحة الكلية الواقعة في أطـراف التـوزيع للمنحنى البياني، والتي قمنـا متعمـدين بتظليلها، واعتبرناها، قيما لمتوسط حسابي غير مألوف للعينة.

نستطيع الآن أن نكون أكثر تحديداً في توضيح توزيع المعاينة الإحصائية، للشكل رقم (11). حيث تفيدنا نظرية النهاية المركزية، بأن للمجتمع البحثى الذي يُقدر متوسطه الحسابي بـ (μ)، وتغارير بـ (σ^2)، إذا ما كانت (n) (حجم العينة) كبيراً نسبيا أو أكثر مـن 30، فـإن توزيـع المعـاينـة الإحصـائية للعيـنـات التـي يقـدر حجمهـا بـ (n) سيكـون ـ تقـريبـاً ـ توزيعـاً طبيعياً، ويكون متوسطها الحسابي مساوياً لـ μ وتغايرها مســاوياً لـ ($\frac{\sigma^2}{n}$)، ولما كان حجم عينتنا يساوي 35، فإن توزيع المعاينة الأحصائية سيكون ـ تقريباً ـ طبيعياً، ولما كان المتوسط الحسابي المفترض لمجتمعنا البحثي يساوي 30، فإن توزيع معاينتنا الإحصائية سوف يكون حوالى 30. وهذه الحقيقة (بأن المتوسط الحســابي أو النقـطة المـركـزية لتوزيع المعاينة الإحصائية للمتوسط الحسابي مساوياً لـ (μ) تثبت قولنا بأن قيمة المتوسط الحسابي ($\overline{X}$)، غير متحيز للمتوسط لحسابي لمجتمع الدراسة الذي تقدر قيمته بـ (μ)، وهـي القيمة الأكثر إحتمالاً لمتوسط حسابي لعينة واحدة، والتي يمكن أن نحصل عليها من إعادة اختبار المعاينة الإحصائية.

نود الآن تحديد تغايرات المتوسط الحسابي للعينة، أو تغاير توزيع المعاينة الإحصائية

للمتـوسط الحسابي ($\bar{X}$). وتشير نظرية النهاية المركزية بأن مجتمع الدراسة الذي يتميز بتغاير يساوي (σ^2)، فإن تغاير توزيع المعاينة الإحصائية ($\bar{X}$) ، لعينات بحجم (n) ، سوف يساوي ($\frac{\sigma^2}{n}$)، وذلك في حالـة معرفتنا ـ على وجـه التحديد ـ للتغاير الحقيقي لمجتمع الدراسة الذي يساوي (σ^2)، ولذلك فنحن مجبرين على تخمينه، وتخميننا الأكثر دقة يعتمد على قيمة تغاير العينة الذي يشار إليه بالرمز (S^2)، وقد قمنا بتعريف (S^2) في الفصل الأول، بالمعادلة التالية:

$$\text{تغاير العينة} = S^2 = \frac{\sum [X - \bar{X}]^2}{n}$$

في المعادلة السابقة، نجد قيمة S^2 على ضوء معطيات الإحصاء الوصفي، بينما نحن الآن بصدد التعامل مع الإحصاء الاستنباطي أو الاستدلالي، ولذا فعلينا أن نُعدل قليلاً من معادلتنا الخاصبة بـ S^2 .

وقد عرفنا، بأن المتوسط الحسابي للعينة، يُعد تقديراً غير متحيز للمتوسط الحسابي لمجتمع الدراسة، وبأسلوب آخر، فإن ($\bar{X}$) تقدير غير متحيز لـ μ ، بينما يُعد تغاير العينة تقديراً متحيزاً لتغاير مجتمع الدراسة، وبأسلوب آخر فإن (S^2) تقدير متحيز لـ (σ^2).

ويُعـد التقـدير متحيزاً، إذا كان غير مساو لمعلمة مجتمع الدراسة التي يسعى إلى تقديرها، والنقطة المركزية (المتوسط الحسابي) لتوزيع المعاينة الإحصائية لتغير العينة S^2 (تغـاير مجتمع الدراسة)، ولكنه أصغر من (σ^2). فإذا أردنا تصحيح هذا التحيز، فعلينا أن نستخدم المعادلة التالية، إذا ما كنا بصدد التعامل مع الإحصاء والاستنباطي:

$$\text{تغاير العينة} = S^2 = \frac{\sum f [X - X]^2}{n-1}$$

وطالما نحن نقسم على مقام أصغر قيمة (1 − n) في هذه المعادلة، فإننا ـ بالضرورة ـ سوف نحصل على تقدير أعلى، وبذلك نكون قد حصلنا على تقدير غير متحيز للتغاير الحقيقي لمجتمع الدراسة.

بالعـودة إلى منـاقشـة تغـاير توزيع المعاينة الإحصائية للمتوسط الحسابي، نستطيع حساب تغاير العينة، باستخدام البيانات التي وردت في جدول (11) كما يلي:

$(\bar{X}-X)^2f$	$(\bar{X}-X)^2$	$(\bar{X}-X)$	$\bar{X}$	F	X
15.735.38	7.867,69	88.7	11.3	2	100
16.90	1,69	1.3	11.3	10	10
595.35	39.69	6.3-	11.3	15	5
266.45	53.29	7.3-	11.3	5	4
383.07	127.69	11.3-	11.3	3	0
16.997.10					

بتعويض المعادلة السابقة، بأرقام هذا الجدول، ينتج الآتي :ـ

تغاير العينة $= S^2 = \dfrac{\Sigma f[\bar{X}-X]^2}{n-1}$

$$S^2 = \frac{16,997,15}{34} = 499,92$$

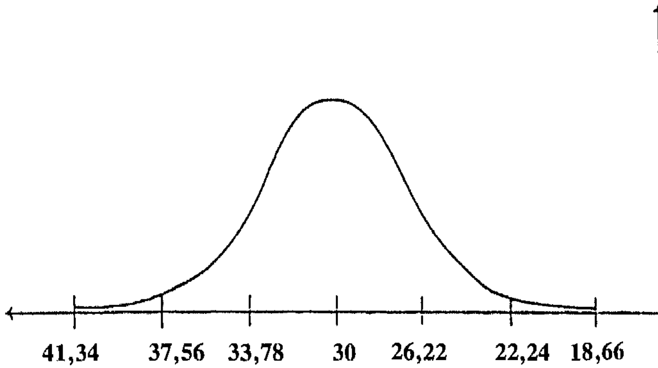
تظهر (f) في المعادلة بسبب وجود أكثر من حالة بحثية لكل قيمة من قيم (X) ، وإذا لم نضـاعف (نضرب) العدد الذي لدينا في عدد مرات تكرار الظهور، لكان علينا أن نستخدم 5 بدلا من 35 لتعويض قيمة (n) . وعلى ضوء هذه المعلومات يمكننا تحديد قيمة تغاير توزيع المعادلة الإحصائية كالآتي :

تقدر قيمته : $\dfrac{\sigma^2}{n} = \dfrac{S^2}{n} = \dfrac{499,93}{35} = 14,28$

وحقيقـة الأمـر، أننـا مهتمـون بتقـدير الانحراف المعياري لتوزيع المعاينة الإحصائية للمتـوسط الحسـابي ($\bar{X}$). والـذي يساوي الجذر التربيعي لقيمة تغايره $\sqrt{\dfrac{\sigma^2}{n}}$ والإشارة إليه بالخـطأ المعياري للمتوسط الحسابي Standard Error ، ويرمز له بالرمز ($\sigma\bar{X}$)، وتقدير قيمته بالمعادلة التالية :

الخطأ المعياري $= \sigma\bar{X} = \sqrt{14,28} = 3,78$

وبـالتنسيق بين هذه المعلومـات، وبمعرفتنا بأن التوزيع المستخدم هو توزيع طبيعي - تقريباً ـ نستطيع أن نقوم، برسم الشكل رقم (12).



11.34 − 30 =	(3.48)3 − 30 =	18.66
7.56 − 30 =	(3.78)2 − 30 =	22.44
	(3.78) − 30 =	26.22
	μ =	30
	(3.78) + 30 =	33.78
7.56 + 30 =	(3.78)2 + 30 =	37.56
11.34 + 30 =	(3.78)3 + 30 =	41.34

شكل (12): توزيع نظري للمعاينة الإحصائية لـ ($\bar{X}$) (المتوسط الحسابي للعينات)، العينات يقدر حجمها بـ 35 ، مُختارة من مجتمع دراسة يقدر متوسطه الحسابي μ بـ 30 ، وتغاير مجتمع بحثي (σ^2) مقداره = 499.92.

يمثل الشكل 12 ، الرسم التوضيحي لتوزيع معاينة إحصائية حيث المتوسط الحسابي للعينة = ($\bar{X}$) لمجتمع دراسة متوسطة الحسابي = 30 ، مستنداً على عينات بأحجام تساوي 35 ، وقيم المتوسطات الحسابية للعينة المسجلة على الخط البياني الأفقي (22.24.18.66 الخ . . . . .) محسوبة كما هو موضح في العمليات الحسابية أسفل الشكل (12) ، حيث 18,66 تبعد عن المتوسط الحسابي لمجتمع الدراسة (30) بثلاثة انحرافات معيارية، وكذلك الحالة بالنسبة لـ 41,34 ، بينما 22,24 ، 37,56 ، تبعد كل منهما عن المتوسط الحسابي لمجتمع الدراسة (30) ، بانحرافين معياريين، ونجد أن 26,22 ، 33,78 تبعد كل منها بانحراف معياري واحد عن المتوسط الحسابي لمجتمع الدراسة (30).

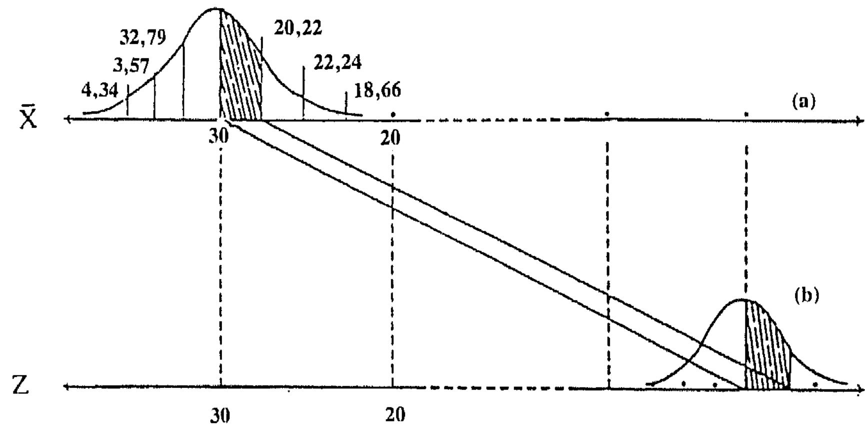
وإذا ما كانت النقاط 16,66 ، 41,34 تبعد كل منها عن المتوسط الحسابي لمجتمع الدراسة بثلاث وحدات انحراف معياري، فهذا يعني أن قيمة الانحراف المعياري لـ 18,66 تساوي ـ 3 [أي درجة الانحراف المعياري = 3,0(3,0 - Z Score)] طالما أنها أقل من المتوسط الحسابي بثلاث وحدات انحراف معياري. وبنفس المفهوم، يمكن القول

بأن قيمة الانحراف المعياري لـ 41,34 تساوي + 3 (أي درجة الانحراف المعياري = + 3(Z Score + 3) ، طالما أنها تقع في نقطة أعلى من المتوسط الحسابي بثلاث وحدات انحراف معياري . ويمكن تحديد درجات الانحراف المعياري ، كالآتي:

Z	$\bar{X}$	إيضاحات
16.88	3.00	3 وحدات إنحراف معياري أسفل المتوسط الحسابي للتوزيع
22.44	2.00	وحدتي انحراف معياري أسفل المتوسط الحسابي للتوزيع
26.22	1.00	وحدة انحراف معياري أسفل المتوسط الحسابي للتوزيع
30.00	000	المتوسط الحسابي للتوزيع (مجتمع البحث)
33.78	1.00	وحدة انحراف معياري أعلى من المتوسط الحسابي للتوزيع
37.56	2.00	وحدتي انحراف معياري أعلى من المتوسط الحسابي للتوزيع
41.34	3.00	ثلاث وحدات انحراف معياري أعلى من المتوسط الحسابي للتوزيع

يوضح الشكل (13) ، التحولات التي تطرأ على درجة الانحراف المعياري (Z Score).

يمثل هذا الرسم ، قيامنا بتحويل التوزيع النظري للمعاينة الإحصائية إلـى وحـدة توزيع طبيعة ، حيث المتوسط الحسابي لمجتمع البحث (µ) = 30 ، والخطـأ المعياري للمتوسط الحسابي لمجتمع البحث ($\bar{X}$ µ = 3,79 ، وهذا يعني تحوله إلى توزيع طبيعي ،



شكـل (13): تحول درجات الانحراف المعياري لتوزيع المعاينة ، حيث المتوسط الحسابي المعياري µ = 30 ، والخطأ المعياري للمتوسط الحسابي المعياري (σ $\bar{X}$ = 3,78)، في وحدة توزيع طبيعي، لمتوسط حسابي = 0 ، انحراف معياري = 1. والمناطق المظللة في التوزيعات أعلاه، تمثل 34% مـن المجموع الكلي للمساحات تحت المنحنى البياني.

حيث المتوسط الحسابي لمجتمع البحث = صفر، والانحراف المعياري = 1 صحيح، وبذلك نغير المسافة بين النقاط المحددة (العينات أو مفردات مجتمع الدراسة) من 3,87 إلى (صفر)، كما نقوم ـ أيضا ـ بتحويل المتوسط الحسابي من 30 إلى (صفر)، أما نسبة المساحة الكلية بين أي نقطتين محددتين في التوزيع (أي بين العينات أو مفرادات البحث) فستظل ثابتة، وذلك لأن أي من هذه التوزيعات، يُعد توزيعاً طبيعياً، وهذه إحدى خصائص التوزيع الطبيعي، كما وضحناها في الفصل الأول.

يمكن تحويل أي نقطة (مركز العينة أو مفرد العينة) في توزيع المعاينة الإحصائية بجدول (12)، إلى درجة معيارية أو قيمة معيارية Z Score، من خلال المعادلة التالية :ـ

$$\text{الدرجة المعيارية} = Z = \frac{\mu - \bar{X}}{\sigma n}$$

هذه المعادلة تتشابه في مفهومها مع المعادلة التي أوردناها في نهاية الفصل الأول، حيث يتم فيها طرح قيمة المتوسط الحسابي ومتوسط المتوسطات الحسابية)، ويقسم الناتج على قيمة الانحراف المعياري. وباستخدام هذه المعادلة نقوم بتحويل كل تسجيلة من مكانها إلى جهة اليسار بمقدار 30 وحدة، كما تقوم بتخفيض قيمة الانحراف المعياري من 3,78 إلى واحد صحيح، أما النسبة المئوية للمساحات الواقعة أسفل المنحنى الطبيعي الواحد، والمحاطة بقيمتين، ستبقى مُعادِلة في قيمتها لأي مساحات أخرى واقعة تحت المنحنى البياني الطبيعي والمحاطة بقيمتين متماثلتين (وهذا يعني أن القيمتين الرئيسيتين هما نفس القيمتين المتحصل عليهما عن طريق التحويل). مثال لذلك:

تتمركز 34% من المساحة الكلية الواقعة أسفل المنحنى (a)، بين القيمتين (26,22) و (30,00) وباستخدام أسلوب التحويل للدرجات المعيارية، نجد أن القيمة (26,22) تطابق 1,00 في المنحنى الجديد (b)، والقيمة 30,00 تطابق 0,00 والمساحة المحصورة بين 1,00 ، 0,00 في المنحنى (b)، تقدر قيمتها بـ 34% من المساحة الكلية الواقعة أسفل المنحنى البياني. ومن الطبيعي، أننا سوف نحتاج لعمل بعض التضريبات الحسابية، لحساب المنطقة الواقعة أسفل المنحنى (a). وعلى أي حال، في الحالة الخاصة للمنحنى (b)، نجد نتائج العمليات الحسابية المتعلقة بالمنحنى الطبيعي المعياري، أو وحدة المنحنى الطبيعي، مُسجلة في جداول، مثل الجدول رقم (2) في ملحق هذا الكتاب.

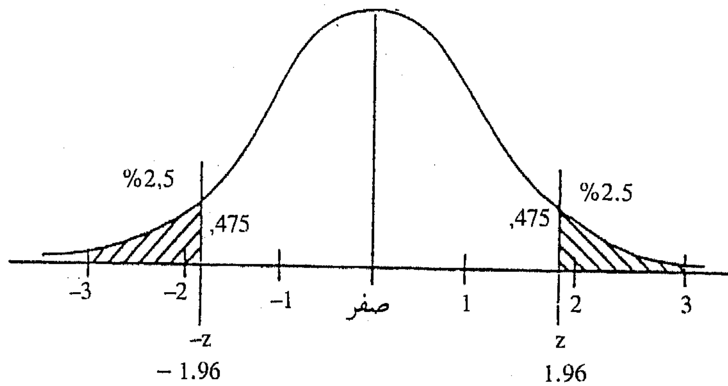
اختبار الفروض المتعلقة بـ μ عندما تكون (n) كبيرة الحجم:

Testirg a hypothesis about μ when n is large

دعنا الآن نرجع إلى بيانات العينة الأصلية، المكونة من 35 مستفيداً. إذ باختيار فروضنا، وجدنا أن المتوسط الحسابي لمجتمع الدراسة مساوياً لـ 30. ونتج عن تحليل بيانات العينة، أن متوسط عدد الكتب المستعارة في السنة الماضية يساوي 11,3 وهنا نتساءل: هل هذه قيمة محتملة أم غير محتملة، لمتوسط حسابي لعينة، إذا ما افترضنا أن المتوسط الحسابي لمجتمع الدراسة يساوي 30؟

أولا، يجب أن نحدد بدقة درجة المخاطرة التي نحن على استعداد لتحملها، برفضنا لفرضية بحث نعلم تماماً أنها صحيحة، ونضع نسبة مئوية لهذه المخاطرة، ونطلق عليها مصـطلح المستوى الدلالي Significance Level ولتكن 5% فإذا وقع المتوسط الحسابي لعينتنـا في المنـطقة المرفوضة، أي في أطراف توزيع المعاينة الإحصائية، حيث توجد المنطقة المظللة التي تمثل نسبة 5% من المجموع الكلي للمتوسطات الحسابية للعينة، سيتحتم علينا الاعتراف بأن نتائجنا غير محتملة الظهور لمجتمع دراسة متوسطة الحسابي 30 وبالتالي سنرفض فرضنا الصفري.

بجانب استخدامنا لتوزيع المعاينة الحقيقي للشكل رقم (12)، فنحن ننوي أيضاً استخدام وحدة التوزيع الطبيعي. ولكن قبل أن نتخذ أي قرار، يجب التأكد من أن الدرجة المعيارية واقعة في بداية المناطق المظللة في الشكل رقم (14).



شكل (14): وحدة منحنى طبيعي، بمناطق مظللة تشير إلى قيم عينة غير محتملة الظهور، والتي تدفعنا إلى رفض فرضنا الصفري.

نقوم بتحويل المتوسط الحسابي لعينتنا من 11,3 إلى درجة معيارية (Z) ، ونرى، ما إذا كانت ستقـع في المنـطقـة المـظللة لشكـل (14) ، فإن فعلت، سنـرفض فرضنا الصفري، الذي ينص على أن المتوسط الحسابي لمجتمع الدراسة يساوي 30.

نقوم بقسمة نسبة الـ 5% (أو 05,) إلى جزئين متساويين، ونضع كل جزء منها في طرف من أطراف الشكل (14) ، أي 025. أو 2.5% في الجانب الأيسر ، و 025. أو 2.5% في الجانب الأيمن. وفي حالة حصولنا على متوسط حسابي صغير جداً أو كبيرة جداً للعينة، سيؤدي ذلك إلى رفضنا للفرض الذي وضعناه، حيث المنطقة الواقعة بين صفر و 2 ، أسفـل المنحنى في شكل (14) ، تساوي 1,457 أو 47,5% من مجموع المساحة الكلية (طالما أن المساحة الكلية فـي الجانب الأيمـن للصفر ، تسـاوي 50, أو 50% مـن المساحة الكلية).

إذا نظرنا في جدول رقم (2) في ملحق الكتاب، وراجعنا القيمة 4750 الواقعة في منتصفه ـ تقريبا ـ سنجد أن قيمة 2 الأفقية (الواقعة على الصف) تساوي 1,9 ، وأن قيمة 2 الرأسية (الواقعة على العمود) تساوي 06, وبجمع هذين الرقمين ينتج لنا درجة معيارية تساوي 1,96 وهذا يعني، أن المنطقة الواقعة بين قيمة الدرجة المعيارية 1,96 ، والمتوسط الحسابي صفر، تساوي 4750, وطالما أن وحدة التوزيع الطبيعي متماثلة، فإن الدرجة الميعارية السالبة، تساوي ـ 1,96 . ولذا، فإن منطقة الرفض (المنطقة المظللة) في الشكل (14) ، التي تدفعنا إلى رفض فرضنا الصفري، تتكون من جميع قيم المتوسط الحسابي للعينة التي تنتج عن درجات معيارية أصغر من 1,96 أو أكبر من + 1,96 ، ويمكن حساب ناتج المتوسط الحسابي للعينة، لهذه الدرجة المعيارية كالآتي:ـ

$$\text{الدرجة المعيارية} = Z = \frac{\mu - \bar{x}}{\sigma\bar{x}} = \frac{30 - 11,3}{3,78} = -4,95$$

بما أن 4,95 تقع في الجانب الأيسر لـ ـ 1,96 ، وداخل منطقة الرفض المظللة، وطالما لدينا مجتمع دراسة متوسطه الحسابي يساوي 30 ، فمن غير المحتمل أن نحصل على عينة متوسطها الحسابي 11,3 ، وبهذا نستنتج أن فرضنا غير صحيح.

الآن قررنا أن افتراضنا: بأن المتوسط الحسابي لمجتمع دراستنا يساوي 30 ، يعد فرضاً خاطئاً (تذكر بأننا قد نكون مخطئين في قرارنا هذا) ونستطيع أن نستنتج بالاستناد على بياناتنا بأن المتوسط الحسابي لمجتمع الدراسة يقدر بأقل من 30% في هذه المسألة، فإن أفضل تقدير للمتوسط الحسابي لمجتمع الدراسة يساوي 11,3.

في مراجعة سريعة، لإختبار الفروض المتعلقة بالمتوسط الحسابي لمجتمع الدراسة، نسجل الخطوات الأساسية التالية، التي يجب اتباعها في هذا الصدد:

1. القيام بعمل تقدير افتراضي لقيمة المتوسط الحسابي لمجتمع الدراسة، وصياغة هذا التقدير في شكل فرض صفري.
2. إذا كان حجم العينة أكبر من 30، فبإمكاننا استخدام وحدة التوزيع الطبيعي لمساعدتنا في اتخاذ القرار، طالما أن نظرية النهاية المركزية تنص على أن توزيع المعاينة الإحصائية للمتوسط الحسابي، يوزع - في هذه الحالة - توزيعاً طبيعياً.
3. اختيار نسبة مخاطرة، أو مستوى دلالي مناسب.
4. إيجاد قيمة الدرجة المعيارية (Z Score)، باستخدام جدول 2 من ملحق الكتاب، الذي يطابق مابين مناطق الرفض المظللة في أطراف بالتوزيع (ونوصي في هذه الحالة، برسم المنحنى البياني للتوزيع وتظليل مناطق الرفض).
5. القيام بعمل عينة عشوائية، ثم حساب المتوسط الحسابي والانحراف المعياري (غير المتحيز) للعينة، وتحديد الخطأ المعياري لتوزيع المعاينة للمتوسط الحسابي للعينات ذات الحجم المطلوب، باستخدام المعادلة التالية :-

$$\text{الخطأ المعياري} = \sigma \bar{x} = \sqrt{\frac{S^2}{n}}$$

6. القيام بحساب قيمـة الدرجة المعيارية (Z) للعينة، باستخدام المعلومات التي حصلنا عليها من العينة (الاختبار الإحصائي)، باستخدام المعادلة الآتية :-

$$\text{الدرجة المعيارية} = Z = \frac{\bar{x} - \mu}{\sigma \bar{x}}$$

7. اتخاذ القرار. لو وقعت قيمة الدرجة المعيارية للعينة داخل منطقة الرفض، نقوم باستبعاد الفرضية الصفرية التي وضعناها. وإذا لم تقع داخل منطقة الرفض، نفترض - مؤقتاً - أن الفرضية الصفرية صحيحة، ونربط مابين النتائج وفرضنا البحثي.

اختبار الفروض المتعلقة بـ (μ) عندما تكون (n) صغيرة:

Testing a Hypothesis about μ When (n) is Small

تنطبق نظرية النهاية المركزية على توزيع المعاينة الإحصائية، للمتوسط الحسابي ($\bar{X}$) عنـدمـا تكون العينة كبيرة نسبياً، أما في حالة كون العينة صغيرة نسبياً، فإن توزيع

المعاينة الإحصائية لـ ($\bar{X}$) لا يتبع التوزيع الطبيعي، بل يتبع ما يسمى بتوزيع (t)، ويشار إليه في العادة بـ توزيع t للطالب "Student's Distribution".

ويُعد توزيع (t)، نوع آخر من الأنواع الخاصة للتوزيعات الرياضية، حيث يتم فيه استخـدام كل معضلات اختبارات الفروض الرياضية، المتعلقة بالمتوسط الحسابي، وذلك عندما تكون العينة صغيرة (عادة أقل من 30). وقد أوردنا سابقاً، أن توزيع المعاينات الإحصائية للمتوسط الحسابي، يكون موزعاً طبيعياً ـ بصفة تقريبية ـ عندما تكون قيمة (n) كبيرة جداً (n> 100). ويصنف توزيع t، كتوزيع طبيعي حقيقي، يُعد استخدامه في اختبار الفروض للمتوسطات الحسابية، عندما تكون n > 100، هو الاستخدام المثالي، حيث يُعد اختبار t هو الاختبار الأكثر دقة، طالما أن توزيع المعاينة الإحصائية للمتوسط الحسابي، يتبع بدقة إجراءات هذا الاختبار (أي خطوات تطبيق اختبار (t):(t Test).

ويختلف شكل توزيع (t) قليلاً، عن الشكل الذي يأخذه التوزيع الطبيعي. ومن الضروري في هذا النـوع من التوزيع أن نضمن عدداً أكبر من وحدات الانحراف المعياري للمتوسط الحسابي، فمثلاً، الأخذ بنسبة 95% من المساحة أو الحالات البحثية الواقعة أسفل المنحنى البياني، وكلما زاد عدد الحالات (n)، كلما بدأ توزيع t أقرب إلى التوزيع الطبيعي، كما سيظهر لنا من فحص جدول t (أنظر الملحق جدول رقم (3).

وكـما توجد عائلة للتوزيع الطبيعي، هناك أيضاً عائلة ذات خصائص مشتركة، لتوزيع (t)، حيث تكـون متوسطاتها الحسابية مسـاوية دائمـاً لصفر، وكـل توزيعات (t) ترتبط بمعلمة يطلق عليها درجة الحرية Degree of Freedom، يشار إليها بالرمز df، وكمـا هـو مألوف فـي جميع تطبيقات توزيـع (t) الرياضية التـي تحتـوي علـى رمز مفرد فإن df = n-1، أي تنقص واحداً عن مجموع عدد الحالات المستقلة، ومفهوم «درجة الحرية» يتصـل بالقيـم التـي تتصف بحرية التغير. ويمكن توضيح درجة ا الحـرية بالمثال الآتي :ـ

لو حددنا متوسط حسابي $\bar{X} = 10$، ونحن على علم بأن هناك 4 حالات في عينتنا، فنحن نملك حرية اختيار القيم الثلاثة الأولى ($X_1 = 5$، $X_2 = 20$، $X_3 = 10$) مما يعطينا مجموع = $5 + 20 + 10 = 35$.

وهذا يعني أن X يجب أن تساوي 5 حتى نتمكن من الحصول على متوسط حسابي يساوي :

$$10 \; (\bar{X} = \frac{5+10+20+5}{4} = 10)$$

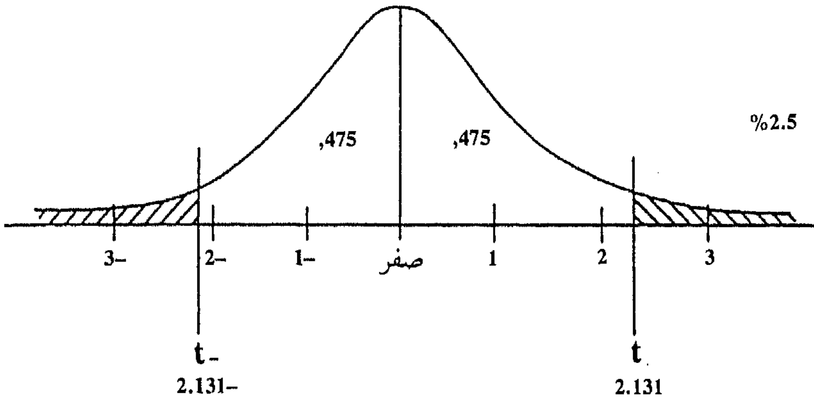
ولهذا، فهناك ثلاث درجات للحرية سُمح لها بالتنوع (أو بحرية التغير). ولإجراء اختبار الفروض لمتوسط حسابي لمجتمع بحثي ما باستخدام توزيع t ، فإن الإجراءات المتبعة تكون مشابهة تماماً لتلك التي ذُكرت في الجزء السابق، مع وجود اختلاف واحد مهم، حيث يمكن تطبيق توزيع (t) على توزيع العينة للمتوسط الحسابي، عندما تكون n > 30 ، بشرط أن يكون توزيع المتغيرات ـ موضع اهتمامنا ـ في مجتمع الدراسة الأصلي ـ الذي شكلت منه العينة ـ توزيعاً طبيعياً، ولا نحتاج لهذا المطلب عندما تكون العينات كبيرة نسبياً، لأن نظرية النهاية المركزية ـ في هذه الحالة ـ يكون لها تأثيرها الإيجابي على التوزيع الطبيعي، ولكن مع العينات الصغيرة فنحن نحتاج لهذا الشرط الإضافي.

وكما هو مألوف، يمكن إيجاد منطقة الرفض لاختبار الفروض للمتوسط العددي، بإدخال المستوى الدلالي أو نسبة المخاطرة ـ التي نحن على استعداد لتحملها ـ على الجدول، ويستخدم لذلك جدول (3) [من ملحق الكتاب]، المتعلق بتوزيع t ويتطلب ذلك معرفة قيمة df (التي تساوي n – 1)، حتى نحصل على النقاط المحددة على حافة منطقة الرفض المظللة، وسنورد مثال يوضح تطابق الإجراءات التي سبق ذكرها مع الإجراءات الحالية.

إذا علمنا أن متوسط سعر الكتب المرجعية هذه السنة بلغ 20 دولاراً، بناء على معلومات صدرت حديثاً في تقرير، وقد وجدنا بعد تحليل عينة عشوائية من الكتب التي تم شراؤها في تلك السنة للمكتبة، بأن المتوسط الحسابي لأسعار الكتب يقدر بـ 24 دولار (X = 24)، وقد قدرنا الخطأ المعياري الذي أخذنا به بـ 3 دولارات (± 3 دولار)، وبما أن العينة العشوائية التي قمنا بتحليلها بلغت 16 كتابا من مجموع الكتب المرجعية المشتراة، (وقد وجدناها عينة معقولة بالنسبة للعدد الكلي للكتب)، فقد كانت فرضيتنا الصفرية، أنه لا يوجد فرق ملحوظ بين متوسط سعر الكتاب في تحليلنا وبين ماجاء في التقرير، وخاصة وأن فرضيتنا الصفرية، قدرت المتوسط الحسابي لمجتمع البحث بـ 20 دولاراً.

ولاختبار فرضنا، يجب علينا افتراض أن توزيع مجتمع الدراسة لأسعار الكتب المرجعية، توزيع طبيعي، وهو كما يبدو افتراض منطقي للغاية، وبسبب حجم هذه العينة، يجب تطبيق اختبار t عليها.

دعنا نفترض أننا وضعنا 5% ، كمستوى دلالي (نسبة مخاطرة)، ثم وجب علينا استخدام جدول 3 [في الملحق]، لإيجاد قيم t في أطراف مناطق الرفض، وكما هو موضح في شكل (15) ، فقد قسمنا الـ 0,05(5%) لجزئين متساويين (كما فعلنا سابقاً في شكل (14) ، وراجعنا جدول (3) في المنطقة التي تساوي 025, وبالنظر عبر الصف



شكل (15): توزيع t ، حيث (1 – n) = (1 – 16) = 15 ، مع منطقة مظللة تشير إلى منطقة الرفض للفرضية الصفرية

الأفقي حيث [df = 15 n – 1 = 16 – 1 = 15] ، والعمود الرأسي حيث يوجد العدد 05, سنجد أن قيمة t = 2,131 ، وبالرغم من أن جدول t يحتوي على القيم الموجبة فقط ، إلا أننـا ، بأخذنا بمبدأ التماثل ، نعلم أن قيمة t في حافة المنطقة المظللة من الجانب الأيسر ، تساوي 2,131 وباستخدام معلوماتنا عن العينة المختارة نستطيع أن نحسب الخطأ المعياري ، للمتوسط الحسابي ، باستخدام المعادلة التالية :ـ

$$\text{الخطأ المعياري} = \sqrt{\frac{S^2}{n}} = \sqrt{\frac{9,00}{16}} = ,75$$

الآن نستطيع حساب إحصاءات اختبار العينة ، والذي سيكون في هذه الحالة قيمة t والمعادلة المستخدمة ، شبيهة بمعادلة درجة الانحراف المعياري (Z).

$$\text{قيمة t للمتوسط الحسابي للعينة} = \frac{\bar{x} - \mu}{\sigma \bar{x}} = \frac{24 - 20}{,75} = \frac{4}{,75} = 5,33$$

هذه القيمة تقع داخل منطقة الرفض ، وهذا يعني أنها بعيدة كل البعد عن فرضية المتوسط الحسابي لمجتمع البحث الذي يقدر بـ 20 دولاراً ، مما يؤدي إلى استنتاج أن فرضيتنا غير صحيحة .

حصلنا الآن على نتائج إحصائية مؤكدة ، مما يجعلنا نتساءل ، ما إذا كانت العينة المستخرجة ـ بالفعل ـ من مجتمع بحثي يبلغ متوسطه الحسابي 20 دولاراً . ومرة أخرى ، لدينا طريقتين رئيسيتين لاختبار صحة فروضنا الصفرية ، فيما يخص المتوسط الحسابي

لمجتمع البحث، أولها التوزيع الطبيعي، وثانيها، توزيع t .

عندما تكون العينة صغيرة للغاية ولدينا شبه قناعة بأن توزيع مجتمع الدراسة طبيعي بدرجة معقولة، يجب علينا في هذه الحالة تطبيق اختبار t . أما إذا كانت العينة كبيرة، وكنا نود الحصول على نتائج دقيقة، فنستطيع - أيضاً - تطبيق اختبار t ، وإن كان تطبيق اختبار t هنا لا يعطي نتائج تختلف كثيراً عن تطبيق اختبار Z ، أما إذا كانت العينة كبيرة نسبياً، فلا فرق إطلاقاً بين تطبيق اختبار t أو اختبار Z ، ولا حاجة بنا لافتراض الشكل الطبيعي في توزيع مجتمع الدراسة، حيث توزيع المعاينة للمتوسط الحسابي في العينات الكبيرة نسبياً، سيصل إلى حد التوزيع الطبيعي في كل الأحوال.

حدود الثقة Confidence Limits

إن أفضـل تقديراتنا لمجتمع البحث، عن طريق تشكيل عينة وحساب متوسطها الحسابي، يتبلور في استخدامنا للعينة الإحصائية. ونعلم أن لجوءنا إلى أسلوب إعادة المعاينة الإحصائية، يكون بغرض التأكيد من مدى صلاحية المتوسط الحسابي للعينة، ويترتب على هذا الأسلوب حصولنا على عدة متوسطات حسابية. وللتأكد من صلاحية هذه المتوسطات الحسابية، نسلك طريق حساب ما يسمى بحد الثقة أو فاصل الثقة Confidence Interval.

تتكون حدود الثقة من نقطتين، إحداهما تعلو المتوسط الحسابي، والأخرى تكون أسلفه، وبتحديدنا لهذه الحدود، نتوقع من المتوسطات الحاسبية المتكررة للعينة أن تقع داخـل هذه الحـدود المرسومة، ويقدر معامل الثقة المعتاد بـ 95,. الذي على ضوئه، نفترض بأننا إذا أعدنا تشكيل العينة 95 مرة من 100 مرة (أي 95%)، فإننا نتوقع أن تقع المتوسطات الحسابية لهذه العينات داخل تلك الحدود أو الفواصل، في كل مرة من الـ 95 (أي بنسبة 95%).

لنفـترض أن لدينـا عينة حيث n = 100 ، ومتوسطها الحسابي $(\bar{X}) = 50$ ، وانحـرافهـا المعياري $(S) = 15$ ، ففي مستـوى 95 نقـوم بمضـاعفة الإنحراف المعيـاري بمقـدار 1,96 ، وعليه فالمنـطقـة الواقعة أسفل وحدة المنحنى الطبيعي، ستكون محاطة بالقيم، + 1,96 و -1,96 وكما تذكر، فهناك 95% من المساحة واقعة تحت هذا المنحنى.

والمعادلة العامة لإيجاد حدود الثقة، كما يلي:

$$\bar{X} \pm 1.96\, \sigma \bar{x} \text{، حيث تقدر } \sigma \bar{x} \text{ بـ } \sqrt{\frac{S^2}{n}}$$

وبتعويض هذه المعادلة بمعطيات مسألتنا الحالية نجد :

$$\bar{X} \pm 1,96\sqrt{\frac{S^2}{n}} = 50 \pm 1,96\sqrt{\frac{(15)^2}{100}} = 50 \pm 2,94 = 47,06 ، 52,94$$

ونتوقع في حالة أخذنا لعدد كبيرة من الحالات ، أن يقع المتوسط الحسابي مابين 47.06 و 52.94 ، وذلك بنسبة 95% من الحالات . أما إذا أردنا استخدام حدود ثقة بنسبة 99% فعلينا استبدال القيمة 2,58 بالقيمة 1,96 . والمسافة التي تقع مابين الحد الأعلى والحد الأدنى ، هي مؤشرنا إلى مدى إمكانية اعتمادنا على المتوسط الحسابي للعينة (أي مدى صلاحية المتوسط الحسابي) ، فمثلاً ، مسافة تقدر بـ 05, تعني أن توزيع المعاينة للمتوسط الحسابي للعينة ضعيف للغاية ، وتعني أيضاً أن المتوسطات الحسابية متلاصقة إلى حد بعيد . أما إذا كانت المسافة 50 ، فهذا يشير إلى أن المتوسطات الحسابية للعينة متباينة إلى حد كبير، وأنها موزعة بدرجة كبيرة حول المتوسط الحسابي لمجتمع الدراسة . وفي حالـة إيجـادنـا لحدود الثقـة لعينـات تتكـون من عدد قليل من الحالات ، يفضـل استخـدام قيـم t بدلاً من استخدام قيم Z كما كنا نفعل مع اختبار الفروض .

اختلاف اختبار المتوسطات الحسابية — Difference of Means Test

يكثر استخدام اختبار t عندما يكون لدينا عينتان عشوائيتان مستقلتان ، تنتميان إلى مجتمع دراسة موزع توزيعاً طبيعياً ، ونبدأ في التساؤل عما إذا كانت الاختلافات مابين المتـوسـطات الحسابية للعينة ، يرجع إلى الصدفة ، أم إلى الاختلافات الإحصائية . وبعبارة أخرى ، هل العينتان مستخرجتان من مجتمعين دراسيين متماثلين في متوسطهما الحسابي؛ أم أنهما مستخرجتان من مجتمعين دراسيين متساويان ، كما تتساوى متغيرات هذين المجتمعـين ، ولنفـترض أن لدينـا عينتين عشوائيتين مستقلتين لكتب في مجال الفيزياء والعلوم الاجتـماعية ، وقمنـا بحسـاب متـوسط أسعـارها ، ونتج لدينا الآتي (لاحظ ، أن لدينا عينات صغيرة) .

العلوم الفزيائيـــة	العلوم الاجتماعية
$\bar{X}_1$ = 17,90 دولاراً	$\bar{X}_2$ = 1, 40 دولاراً
S_1 = 2.30	S_2 = 2 , 10
n_1 = 20	n_2 = 30

وسؤالنـا هو : هـل هناك فـرق جـذري بين أسعـار كتب الفيزياء وأسعار كتب العلوم الاجتماعية ؟

للإجابة على هذا التساؤل، نضع افتراضاً بأنهما لا يختلفان، وأن المتوسط الحسابي للأسعار لمجتمع الدراسة غير مختلفة (وهذا هو افتراضنا الصفري)، ويمكننا حساب الفرق بين المتوسطين الحسابيين لهذا المثال، وسنجد أن قيمتهما لن تساوي صفراً. ولو كان المتوسطان الحسابيان متساويين في القيمة، فإن الفرق بينهما سيساوي صفراً. وفي حالة ما إذا كان المتوسطان الحسابيان لمجتمع الدراسة متساويين، وإننا أخذنا عينتين عشوائيتين مستقلتين من كلا مجتمعي الدراسة (واحده من كل مجتمع)، فإننا سنحصل في بعض الأحيان على فرق إيجابي بين المتوسطات الحسابية للعينة، وأحياناً أخرى على فرق سلبي بين المتوسطات الحسابية للعينة، وغالباً، لن نحصل على أي فروقات بين المتوسطات الحسابية للعينة.

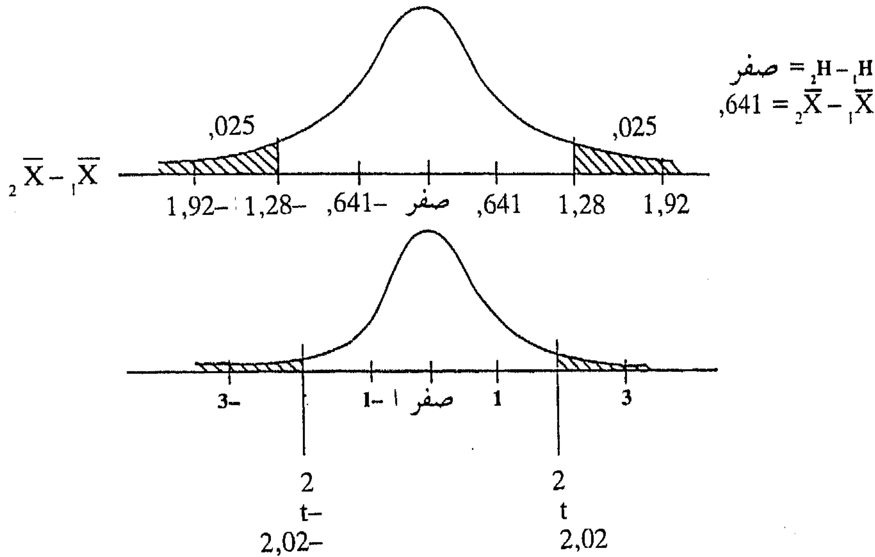
وإذا أخذنا أكبر عدد ممكن من العينات الثنائية من المجتمعين الدراسيين وحددنا الفـروقـات مابين المتوسطات الحسابية. فإنه يمكننا تشكيل توزيع لهذه الفروقات، وسيكـون توزيعـاً للمعاينة الإحصائية العشوائية للفروقات بين المتوسطين الحسابيين للعينـــة. فإذا ماكـــانت الفـرضية الصفـرية صحيحـة، فسيكـون توزيع المعـاينـة الإحصائية كالآتي:

بإستخـدام اختبار t حيث $n_1 + n_2 - 2$ درجة حرية، ومتوسط حسابي يساوي صفراً، وخطأ معياري (للفرق مابين المتوسطات الحسابية) $\sigma_{\bar{X}_1 - \bar{X}_2}$، يمكن استخدام المعادلة التالية:

$$\sqrt{\frac{{}^{1}S}{{}_{1}n} - \frac{{}_{1}S}{{}_{2}n}} = {}_{1}\bar{X} - \sigma\, {}_{1}\bar{X}$$

وبإختيار 05, للمستوى الدلالي، يمكننا تحديد منطقة الرفض، ـ كما تعودنا ـ بالنظر في جدول رقم (3) بالملحق، وبمراجعة الرقم المطابق لهذه المنطقة، سنجد أن أقرب قيمة مسجلة لـ t ، على ضوء المستـوى الذي قيمته، درجة الحرية التي تقدر بـ 48 ، يقدر قيمتهـا بـ 2,02 + 2,02 ، وسـوف نحول فروقات العينة إلى قيمة t فإذا كانت أقل من − 2,02 ، أو أكثر من + 2,02 ، فسوف نستبعد فرضنا الصفري.

يحتـوي شكل (16) على توزيع معاينة إحصائية عشوائية، للفرق مابين المتوسطات الحسابية، مستنداً على بيانات العينة، وعلى توزيع t المطابق مع مناطق الرفض المظللة المناسبة.



شكل (16): توزيع معاينة إحصائية للفرق مابين المتوسطات الحسابية للعينة، مستنداً على حجم العينات: $n_1 = 20$ ، $n_2 = 30$ ، توزيع مطابق لـ t.

ولتحويل فرق العينة لقيمة t ، نستخدم المعادلة التالية:

$$t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{\sigma_{\bar{x}_1 - \bar{x}_2}}$$

مرة أخرى نجد لدينا قيمة ، سنطرح منها المتوسط الحسابي لمجتمعها الدراسي (وبما أن المتوسطات الحسابية لمجتمع الدراسة متساوية، فإن الفروقات بينهما ستكون صفراً) ويقسم الناتج على الخطاء المعياري أو على الانحراف المعياري لتوزيع المعاينة الإحصائية، وتنتج قيمة t للعينة كالآتي:

$$t = \frac{17,90 - 16,40 - 0}{\sqrt{\frac{(2,30)}{20} + \frac{(2,10)}{30}}} = \frac{1,5}{,265 + ,147} = \frac{1,5}{,647} = 2,34$$

ونتيجة للتحليل الرياضي، نستبعد الفرض الصفري، ونستنتج أن الأسعار مختلفة تماماً، إذا ما استخدمنا المستوى الدلالي 05, (نسبة المخاطرة). ونتوقع لنتائج متوسعة (عامة) مثل نتائجنا، بأن لا تظهر بأكثر من احتمال 5% ويلاحظ أن قيمة t التي قمنا بحسابها، لن تكون ذات دلالة في مستوى 01, إذا ما كنا قد قمنا باختياره كمستوى للتحليل.

تشـير نتـائج البحث، بأن لا نتوقع ظهور أي من مفردات العينة في التوزيع إلا بالصدفة البحتة وبنسبة تتراوح مابين 1% و 5%.

ولنأخذ مثالاً آخر، لكيفية إجراء اختبار للفروقات بين المتوسطات الحسابية. إذا كنا بصـدد تحديد مرتبـات ذات مستـويات مختلفة لمكتبيين مبتدئين من الجنسين، وعلى افـتراض أن مرتبات الرجال أكثر من النساء، ونتائج دراستنا لعينة المرتبات المدفوعة خلال الخمس سنوات السابقة تتلخص في الآتي:

ذكـــور	أنـــــاث
$\bar{X}_1 = 10,600$ دولاراً	$\bar{X}_2 = 9,800$ دولاراً
$S_1 = 1,200$	$S_2 = 1,300$
$n_1 = 25$ رجلاً	$n_2 = 36$ إمرأة

وتـظهر نتائج الدراسة لعينة مكونة من 25 رجلاً و 36 إمرأة، أن متوسط ما يدفع للرجل يساوي 10,600 دولار، وما يدفع للمرأة 9,800 دولار، والانحراف المعياري لكل فئة كان 1,200 للرجال، 1,300 للنساء. وعلى ضوء هذه البيانات من الممكن صياغة الفروض التالية:

1. لقد قمنا بتشكيل عينات عشوائية مستقلة.
2. وزعت مجتمعات البحث التي تخيرنا منها العينات، توزيعاً طبيعياً، وتتساوى - أيضاً - في التغاير.

تم قياس التغاير موضع البحث، فيما يخص قيمة المتوسطات الحسابية، على ضوء المستوى الفاصل، وكان افتراضنا الصفري، بأنه لا يوجد فروقات تذكر بين مرتبات الـرجـال والنسـاء. ولذلك افترضنا أن المتوسط الحسابي لمجتمع الدراسة (المرتبات) للرجال مساوٍ لمثيله عند النساء. وقمنا بتحديد قيمة المستوى الدلالي بـ 05, (أي 5% نسبة مخاطرة)، وهذا يعني أننا نخاطر بأن نكون خطأ في فرض افتراضنا الصفري بنسبة 5% ، إذا ما أخـذنـا في الاعتبـار عينـات كبـيرة العـدد. يتم معالجة هذه البيانات - السابقة - حسابياً كالآتي:

$$t = \frac{(X_1 - X_2) - 0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$= \frac{800}{\sqrt{\frac{(1,300}{36} + \frac{(1,200}{25}}}$$

$$= \frac{800}{\sqrt{46,944 + 57,600}}$$

$$= \frac{800}{323,33}$$

$$= 2.47$$

df = درجة الحرية = $n + n - 2$

$= 36 + 25 - 2$

$= 61 - 2$

$= 59$

بمراجعة جدول t ، نجد أن النتائج التي تم التوصل إليها ذات دلالة عند المستوى الدلالي 05, ، مما يعني أننا سنستطيع رفض فرضيتنا الصفرية، التي تنص على أنه لا يوجد أي فروقات بين مرتبات المكتبيين من الرجال أو النساء. وهذا يعني، إذا كانت المتوسطات الحسابية بين مرتبات مجتمع البحث متساوية بين الرجال والنساء؛ كان من المحتمل أن نجد هذا الفرق الكبير كالذي وجدناه بالفعل في 5% من العينات الممكن اختيارها، ولذلك يمكننا استنتاج أن المتوسط الحسابي لمرتبات مجتمع الدراسة للرجال أكبر من نظيره للنساء. ومرة أخرى قد نكون مخطئين في قرارنا، والطريق الوحيد للتأكد من صحة القرار، هو إجراء المزيد من الاختبارات لمعرفة صحة أو خطأ القرار المتخذ. تعرضنا في هذا الفصل لشرح القواعد الرئيسية لاختبارات الفروض، وحصرنا أنفسنا في الإختبارات المتعلقة بالمتوسط الحسابي لمجتمعات الدراسة، وأنواع الاختبارات التي تعرضنا لها مثل اختبار (Z) واختبار (t) ، من الإختبارات الصالحة للتطبيق، في حالة تحليلنا للمتوسط الحسابي لعينة أو إثنتين. تعرضنا ـ أيضاً ـ لبعض اختبارات الفروض الإحصائية المتعلقة بمعلمات مجتمعات الدراسة، بدلاً من المتوسطات الحسابية. ولكن التوزيعات الرياضية التي استخدمت في هذا الصدد، في توزيع المعاينات الإحصائية، قد لا تتصف بأنها توزيعات طبيعية أو توزيعات t (شبه الطبيعية).

وعـلى أي حال، هذه الاختبـارات شائعـة ومعـروفة، ومطبقة، وقد رأينا أنه من

الأفضل التعرض لها وشرحها في هذا الفصل. ومراجع القراءة في نهاية هذا الفصل تغـطي الأنـواع الأخـرى من الاختبـارات المتعلقـة بالمتـوسـطات الحسـابية للعينة، والفروقات بين المتوسطات الحسابية، والحالات الأخرى التي لم نتمكن من تغطيتها في عملنا هذا.

والآن نركز اهتمامنا على العلاقات بين المتغيرات في مجتمع الدراسة.

تمرينات الفصل الثالث:

يناقش بعض العـاملين في المكتبـة، بأن هناك متغيرات قد حدثت خلال السنة الماضية فيما يتعلق بالمرتبات. ومرة أخرى، نأخذ عينات من بدايات المرتبات للرجال والنساء، للسنة الماضية.

وكانت نتائج العينة كالآتي، باستخدام اختبار t ماذا نستطيع أن نقـول، عن هذه المتغيرات المزعومة؟

رجال	نساء
$\bar{X}$ للمرتبات تقدر بـ 10,750 دولار	$\bar{X}$ للمرتبات تقدر بـ 10,491 دولار
S_1 = 150 دولار	S_2 = 200 دولار
n_1 = 5	n_2 = 5

القراءات والمراجع

Blalock, Hubert M., Jr. *Social Statistics*. 2d ed. New York: McGraw-Hill, 1972, p. 188-193, 219-228.

Glass, Gene, and Julian C. Stanley. *Statisitical Methods in Education and Psychology*. Englewood Cliffs, N.J.: Prentice. Hall, 1970.

Hays, William. *Statistics for the Social Sciences*. New York: Holt, Rinehart and Winston, 1973.

الفصل الرابع

الارتباط والانحدار CORRELATION AND REGRESSION

الفصل الرابع
الارتباط والانحدار
CORRELATION AND REGRESSION

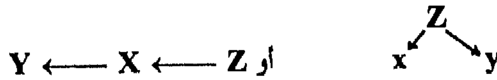
رأينا في الفصول السابقة، كيف تدلنا الاختبارات الإحصائية الدلالية على ما إذا كانت هنـاك علاقـة بين المتغيرات. ومثال لذلك، توصلنا في الفصل الثالث إلى أنه بالاعتماد على عينتين عشوائتين مستقلتين، تمكنا من اكتشاف احتمال وجود علاقة بين نوع الجنس للمكتبين (رجل / أو إمرأة)، والمرتب السنوي الذي يحصل عليه كل نوع، واتضح من تحليلنا للعينة، أن المكتبي الرجل يحصل على مرتب سنوي أعلى بدرجة ملحوظة مما تحصل عليه زميلته في المهنة.

نهتم في المرحلة الحالية بمعرفة مدى قوة العلاقات بين متغيرين، وكيفية التنبؤ بمثل هذه العـلاقـة، ويمكننـا الحصـول على هذه المعلومـات عن طريق معيـاري مهم، باستخدام إجراءات تحليل الارتباط والانحدار.

الارتبـــاط Correlation

الإجـراء الـرئيسي الـذي يؤخذ في الاعتبار في هذا الصدد، يعرف باسم معامل الارتباط الخطي لبيرسون Pearson's Product-Moment Correlation ، ويعرف أيضاً بمعـامل ارتباط الترتيب الصفري Zero-Order Correlation ويرمز إليه بالرمز (r) ، ويعبر عنه برقم عشوائي، تتراوح قيمته مابين (-1 و + 1) ، وتشير أي قيمة تعطى لـ (r) إلى قوة الارتباط الخطي بين المتغيرات، وتشير القيمة الموجبة أو السالبة الكبيرة، إلى أن العلاقة بين المتغيرات قوية، ويوصف الارتباط بأنه موجب (طردي)، عندما تزداد قيمة المتغـير، بإزدياد قيمة المتغير الآخر، ويوصف بأنه سلبي (عكسي) عندما تزداد قيمة المتغير بتناقص قيمة المتغير الآخر، وبدون دليل تمدنا به إجراءات تحليل الارتباط، لا يمكننا اعتبار هذا الارتباط «عرضي».

ومثـال ذلك : المتغير المستقل (X) ، وليكن ميزانية المكتبة، مرتبط بشكل طبيعي وبقوة، مع المتغير (y) (التي تمثل حجم مقتنيات المكتبة)، ويمكن القول بوجود صلة «عـرضية» بين هذين المتغـيرين، بمنـطق، أنـه كلما زادت الاعتـمادات المالية لدى المكتبيين، كلما توسعوا في سياسة شراء المقتنيات، أي (x ←– y) ، هذا المنطق يبدو معقـولاً ومقبولاً، ولكن الصلات العرضية لا يُعتد بها كدليل على الصعيد الرسمي العلمي، فقد يُرد على هذا الادعاء، بأن هذين المتغيرين غير مستقلين أي (x ←– y) ، وأن هناك متغيراً أو أكثر (Z) يتعلق باحتياجات المستفيدين، والرغبة في تلبية هذه الاحتياجات، وهذا هو الذي يحدد حجم الميزانية والمقتنيات، أو يكون السبب المؤثر في تحديد حجمها.



يمكن اعتبار الارتباط دالة للانحدار، حيث يُعد الانحدار ـ بصفة أساسية ـ طريقة خاصة لوضع قانون بالمفهوم العلمي، إذ أنه يمدنا بقاعدة للتنبؤ بقيمة المتغير (يرمز لها بـ S) بالاستناد على قيمة متغير آخر (S) ، والمعادلة التالية هي الشكل الرياضي لتحليل الانحدار المتعلق بمتغيرين :

$$bx + a = y$$

ونحـاول عن طريق إيجـاد قيمة انحدار متغيرين، التنبؤ بقيم متغير واحد (y) ، بالاستناد على قيم المتغير الآخر (X) ، ونحاول الوصول بقدر الإمكان إلى إيجاد قانون «علاقة» بين المتغيرين، ولكن لن نستطيع تجنب ظهور أخطاء في تنبؤنا بقيم (y). ولذا فالعلاقة التي سنوجدها لن تتصف ـ بأي حال ـ بالكمال.

ولذلك، فالمعادلة السابقة، يجب تعديلها لتبدو على الشكل التالي :

$$\text{خطأ} \pm bx + a = y$$

وحقيقة الأمر، أن هذه المعادلة تنص على أنه إذا ما تم التوصل إلى حل لـ a ، b ، يمكننا أن نتنبأ بقيمة المتغير (y) ، بالنسبة لأي قيمة تُعطي للمتغير (x). عادة يكون تحليل الانحـــدار ذو طبيعـــة متعـــددة التغـير، ولهـذا، فإننـا نسعى إلى التنبؤ بقيم المتغير (y) ، عن طريق القيم المتعددة للمتغير (x).

ولنفهم طبيعة الانحدار، من الأفضل أن نبدأ بتحليل انحدار بسيط.

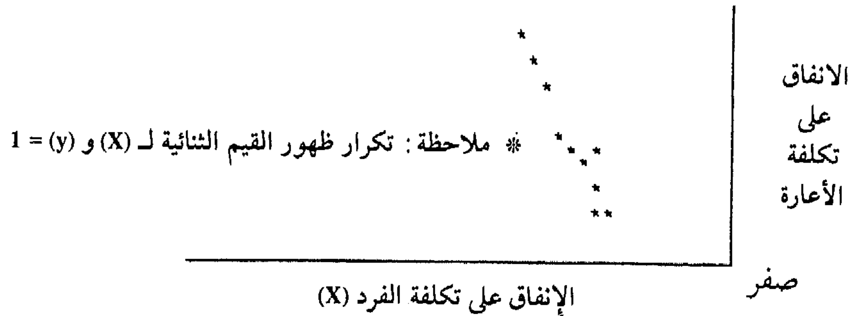
شكـل الانتشـار (الشكل البياني للإنتشار) Scattergrams

يعيننا شكل الانتشار على معرفة ما إذا كان أحد المتغيرات مرتبط بمتغير آخر، وفي بهذه الحـالة نقوم برسم «شكل انتشار»، بحيث توضع كل حالة في مركز معين على الشكل، بناء على قيمتها المأخوذة من المتغيرين (x) أو (y). وشكل الانتشار، عبارة عن رسم بياني ثنائي البُعد، يتكون من نقاط منسقة بناء على قيم المتغيرين موضع الدراسة.

يمثل شكل (17) ، شكـل انتشار يعتمد فـي معطياته على جدول (12) ، الـذي يُظهـر نمط الارتباط بين الإنفاق على تكلفة الإعارة، لعينة مكونة من 10 مكتبات. يقيس الخـط البياني الأفقي (الإحداثي السيني) قيمة الإنفاق على تكلفة الفرد، أما الخط البياني الرأسي (الإحداثي الصادي)، فيقيس قيمة الإنفاق على تكلفة الإعارة. فمثلاً: مكتبة (A) ، يقـدر الإنفـاق علـى تكلفة الفرد بها بـ 4,6 دولاراً والإنفاق على تكلفة الإعارة بـ 4,5 كتاباً.

ويُعـد شكـل الانتشـار، مهـماً للغـاية، لمساعدتنا فـي معرفة طبيعة العلاقة بين المتغيرات، فهو يظهر اتجاهها، وما إذا كانت إيجابية أو سلبية، وأيضاً، ما إذا كانت علاقة خطية.

نجد في الشكل (17) أن العلاقة إيجابية (طردية)، حيث يزداد الانفاق علـى تكلفة الفرد، وكذلك الحال على الإنفاق على تكلفة الإعارة، بالإضافة إلى أنه بإمكاننا رسم خـط مستقيم يمـر ـ تقريباً ـ بالنقـاط كلهـا ، وهـذا يعنـي وجـود علاقـة (ارتباط) خطية، والعلاقة بين المتغيرات (y) و (x) علاقة خطية واحـدة، حيث يعتـمد (y) رياضيا على (x). بمعنى أن (y) تعتمـد علـى (x) ، وليس على X_1 أو X_2 . . . الخ . ولو جمعنا



شكل (17): شكل انتشار يحتوي على نقط بيانات لـ 10 مكتبات، بالنسبة إلى متغيرين (x) و (y)

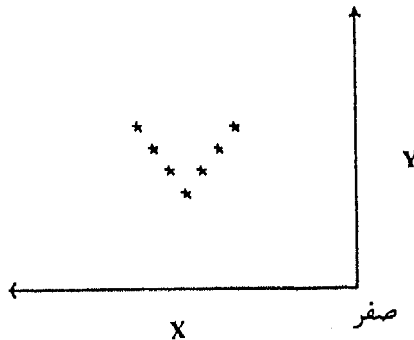
جدول (12): بيانات 10 مكتبات مقاسة على ضوء متغيرين، الإنفاق على تكلفة الفرد (x) ، الإنفاق على تكلفة الإعارة (y)

الحـالات	المتغير x	المتغير y
A	4,60 دولار	4,5 كتاب
B	4,10	4,6
C	6,70	8,2
D	3,90	2,5
E	3,00	2,1
F	5,10	4,9
G	4,00	3,9
H	7,2	8,9
I	6,20	7,7
J	5,50	5,4

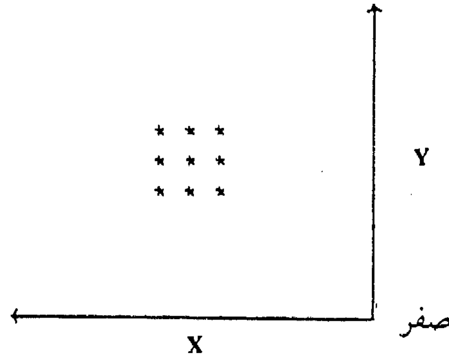
قيم $\frac{n}{r}$ لـ (X) و (Y) ، سيتطابق مع المعادلة :

$bX + a = y$. وستكون النتائج دائماً خطاً مستقيماً، بالرغم مـن أن رسمـنا البياني يظهر خطوطاً مختلفة تعتمد على قيم a و b التـي سنستخدمها وعمومـاً، تتطلب الاستفادة مـن الارتباط الخطي أن نفترض ـ بثقة ـ وجود علاقة خطية .

يظهـر شكل (18) ، علاقة جديرة بالاهتمام، حيث خطها البياني «غير خطي» ، وإجراء (r) غير مناسب للتطبيق، وقد يظهر شكل الانتشار ـ أيضاً ـ أنه لا يوجد ارتباط



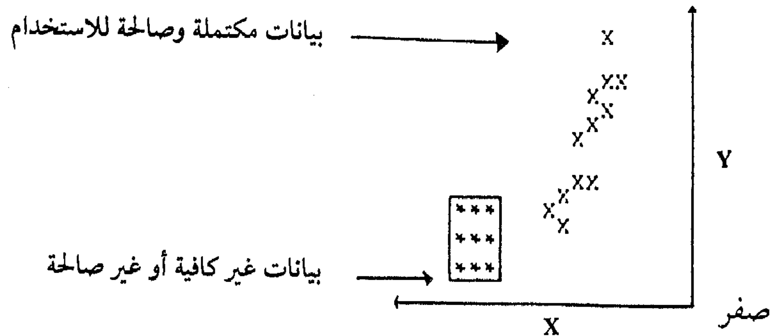
شكـل (18): شكـل انتشـار، يظهـر العـلاقـة بين متغـيرين، حيث معـامل بيرسون (r) ، ليس بالإحصاءات المناسبة للاستخدام



شكل (19): شكل انتشار يوضح عدم ارتباط بين المتغيرات (x) و (y)

خطي أو غير خطي هام، وهذا ما يوضحه شكل (19) ، حيث توجد قيم متعددة لـ (y) مقابل لـ (X) ، والعكس صحيح .

ويستعـرض شكـل الانتشار (20) ، مشكلة قد تطرأ عندما لا توفر الدراسة عينة (بيانات) صالحة أو كافية، كتلك البيانات المسجلة داخل المنطقة المحصورة بالمستطيل على الـرسم البياني، فلو احتـوت الدراسة على بيانات أكثر، أو على عينة أكثر تمثيلاً لمجتمع الدراسة، لاحتوت على الحالات المبينة خارج إطار المستطيل، وفي هذه الحالة، سنفـترض ـ خطأ ـ وجـود ارتبـاط واهي للغـاية، أو عدم وجـود ارتبـاط مطلقاً بين المتغيرات، وذلك استناداً على البيانات التي زودتنا بها الدراسة. في حين أن دراسات أخـرى لاحقـة، حول نفس مجتمـع الـدراسة، قد تسفر عن وجود ارتباط مهم بين المتغيرات نفسها.



شكل (20): شكل انتشار، يوضح عينة غير ممثلة لمجتمع الدراسة، مما ينتج عنها، تقديرات واهية للارتباط بين المتغيرين (x) ، (y)

عندما تكون العينات قليلة ـ نسبيا ـ يكون من السهل معرفة إذا ماكان هناك نوع من العـلاقـة، وأيضـا معرفة، ما إذا كان هناك متغير مرتبط مع متغير آخر. فعندما نفحص أعمدة الأرقام لـ (x) و (y) في جدول (12) ، نلاحظ أنه مع ازدياد نفقات تكلفة الفرد، تزداد نفقات تكلفة الإعارة أيضاً.

وتنبـع أهمية شكـل الانتشار من قدرته على التلخيص الدقيق للبيانات في صورة واحدة بسيطة، ولذلك فإنه من الصعب إستخدام هذا الشكل في تخمين اتجاه مسار البيانات بدقة (يعني صعوبة تحديد الارتباط الخطي بين المتغيرات موضع البحث)، عنـدما يكون هناك العديد من الحالات البحثية، وبالرغم من إمكانية رسم «شكل الانتشار»، يدوياً بسهولة على ورق رسم بياني، إلا أننا ـ غالباً ـ ما نحصل على البيانات في شكل مقروء آلياً (أي عن طريق الحاسوب)، وإنتاج الرسومات بالحاسوب، أصبح أمـراً في غاية السهـولـة. ومثـال ذلـك البرامج الإحصائية، كبرنامج (SPSS) للعلوم الاجتماعية "Statistical Package for Social Sciences" ، ومما لاشك فيه أن الحصول على نسخة مطبوعة مستخرجة من الحاسوب لشكل الانتشار، ومسجل عليها البيانات والإحصاءات الرقمية (r) ، ستعيننا على تفسير وتحليل النتائج بصورة أفضل.

حساب معامل الارتباط (r) — Calculating (r)

تتميز خطوات حسـاب معامل الارتباط (r) ، بأنها مباشرة وغير معقدة. وكما هو موضـح في المثال التالي، فإن إجراء العمليات الحسابية لحالات كبيرة العدد، يكون مرهقاً بدون استخدام إمكانات الحاسوب. وبالتوافر الحالي للحواسيب الآلية، فمن الأفضل استخدامها في هذا المجال إذ لها القدرة الفائقة على استخراج نتائج جاهزة، كما بإمكانها ـ أيضاً ـ عمل رسوم لأشكال الانتشار، وتستخدم المعادلة التالية لحسـاب معـامل الإرتباط (r):

$$r = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2 - (\Sigma x)^2]\,[n\Sigma y^2 - (\Sigma y)^2]}}$$

مثال: ماهو الارتباط بين نفقات تكلفة الفرد (x) ، ونفقات تكلفة الإعارة (y) ، بين الـ 10 مكتبات (n) ، في جدول (13) ؟

بالتعويض في المعادلة السابقة:

$$r = \frac{n\,\Sigma y - (\Sigma x)(\Sigma y)}{\sqrt{[\,n\,\Sigma x^2 - (\Sigma x)^2\,][\,n\,\Sigma y^2 - (\Sigma y)^2\,]}}$$

$$= \frac{10\,(292,66) - (50,30 \times 50,70)}{\sqrt{[\,10\,(269,61) - (50,30)^2\,][\,10\,(326,19) - (52,7)^2\,]}}$$

$$= \frac{2,926,60 - 2,650,81}{\sqrt{(2,696,10 - 2,530,09)\,(3,261,90 - 2,777,29)}}$$

$$= \frac{275,79}{\sqrt{(166,01)\,(484,61)}}$$

$$= \frac{275,79}{\sqrt{80,450,106}} = \frac{275,79}{283,64} = ,972 = ,97$$

جدول رقم (13): بيانات مجمعة من 10 مكتبات لمتغيرين :
نفقات تكلفة الفرد (x) ، نفقات تكلفة الإعارة (y)

المكتبة	الاتفاق على تكلفة الفرد	X^2	الاتفاق على تكلفة الإعارة	Y^2	XY
A	4,60 دولار	21,16	4,5	20,25	20,70
B	4,10	16,81	4,6	21,16	18,86
C	6,70	44,89	8,2	67,24	54,94
D	3,90	15,21	2,5	6,25	9,75
E	3,00	9,00	2,0	4,41	6,30
F	5,10	26,01	4,9	24,01	24,99
G	4,00	16,00	3,9	,2115	15,60
H	7,20	51,14	8,9	79,21	46,08
I	6,20	38,44	7,7	59,29	47,74
J	5,50	30,25	5,4	29,16	29,70
المجموع	50,30	269,61	52,7	326,19	292,66

$\Sigma X = 50,30$

$\Sigma X^2 = 269,61$

$\Sigma y = 52,7$

$\Sigma y^2 = 326,19$

$\Sigma xy = 292,66$

بما أن قيمة (r) تتراوح مابين −1 و+ 1 ، وأن ا لنتائج أظهرت أن قيمة (r) = 0,97 فإن هناك ارتباطاً إيجابياً قوياً (طردياً) بين المتغيرات . وأن مقدار التغير في قيمة (y) مرتبط بقوة مع مقدار التغير في قيمة (x) ، وبدقة أكثر فإن أي زيادة في قيمة (x) ، تصاحبها زيادة متكافئة في (y) ، ولذا فإن الارتباط ـ هنا ـ إيجابي (طردي) .

وينبغي أن نتـذكـر أنه بإمكاننا ـ في حالات أخرى ـ الحصول على ارتباط سلبي (عكسي) قوي ، ويعني في هذه الحـالـة ، بأن الـزيادة في قيمة (x) سيصاحبها نقص مكـافيء في قيمـة (y) . فمثـلاً : بين مجتمع البالغين المستفيدين من المكتبة ، يمكننا الحصول على مثل هذا الارتباط السلبي ، حيث يقل استخدام المكتبة بين البالغين كلما تقدم بهم العمر (أي عندما يزداون سناً) ، أي ترتبط أعمارهـم بعلاقة سلبية (عكسية) مع استخدامهم للمكتبة .

ومن المهم ، أن نتـذكر دائماً ، أن الارتباط ، واتجاه الارتباط (طردي أو عكسي) . وعندما يكون معامل الارتباط (r) مساوياً لـ 87, يوصف الارتباط بأنه أكثر أهمية ، عما إذا كان معامل الارتباط (r) مساياً لـ 32, وأن العلاقة السلبية (العكسية) ، توضح لنا الصورة ، بنفس القوة التي توضحها العلاقة الإيجابية (الطردية) . وتظهر قوة الارتباط عن طريق حجم وقيمة معامل الارتباط (r) ، وليس عن طريق مجرد وجوده .

اختبار الفروض حول ارتباط مجتمع الدراسة :

Hypothesis Tests About the Population Correlation

بما أن البيانات تستخرج من عينة ـ وغالباً ما تكون العينة صغيرة ـ فإنه من الأفضل أن يُختبر الافتراض الصفري ، أي أن الارتباط بين المتغيرين لمجتمع الدراسة يساوي صفراً . يُرمز لمعامل الارتباط الخطي لبيرسون في مجتمع الدراسة بالرمز P(rho), ولذا فنحن نختبر الفرضية القائلة بأن P = صفر .

يمكننا أن نبدأ باختيار مستوى دلالي (نسبة مخاطرة) ونختبره ، لنرى ما إذا كانت قيمة العينة (r) ، تُعد قيمة غير محتمل الحصول عليها ، عند غياب الارتباط الخطي بين

متغيرات مجمع الدراسة . وكما يحدث في اختبارات الفروض للمتوسطات الحسابية ، لو كان الارتباط الحقيقي لمجتمع البحث يساوي صفراً ، فعلينا أن نراجع توزيع المعاينة الإحصائية لارتباط العينة (r). وهذا يعني ضرورة تحديدنا ، لما سوف تعطينا له قيم عينة (r) من نتائج ، لتدفعنا إلى رفض فرضنا الصفري . ولحسن الحظ ، فإن إجراءات هذا الاختبار أسهل بكثير من إجراءات اختبارات الفروض للمتوسطات الحسابية . وذلك ، بسبب وجـود جداول جاهزة التحضير، لقيم معامل الارتباط ، للمستويات الدلالية المختلفة ، ولأحجام العينات (نجدها في جدول (5) بالملحق) . ولاستخدام الجدول ، نحسب أولاً درجـة الحرية (2 – n = df) ، ثم نستخرج قيمة (df) على ضوء المستوى الدلالي المختار (من جدول (5) في الملحق) .

ولتطبيق ذلك على مثالنا : 8 = df ، المستوى الدلالي = 05, ووجدنا في الجدول رقم (5) ، أن الحـد الأدنى للارتبـاط المتوقع لرفض الفرضية الصفرية يساوي 65, ، إذا ماكانت درجة الحرية (df) = 8 ، بمعنى ، إذا ماكانت القيمة الجدولية (r) ، أقل من قيمة (r) التي حصلنا عليها ، يمكننا افتراض أن قيمة معامل ترابطنا (r) ، ذات دلالة إحصائية ، بعيدة بما فيه الكفاية عن قيمة الصفر، لتجعلنا نقرر بأن افتراضنا الصفري خاطيء ، ويعني هذا وجـود ارتباط خطي . ولكن ، هذا لا يعني بأن هناك أي صلة عرضية بين المتغيرات ، بل قد يعني أن المقياس الذي وجدناه ، لا يرجع فقط إلى محض الصدفة ، وقد يكون هناك عامل آخر ـ لم يقاس ـ هو المسبب لهذه العلاقة .

تحليل الانحـــدار **Regression Analysis**

عندما يكون هناك ، ارتباط قوي بين المتغيرات ، كما في المثال السابق ، فإننا نحتاج إلى التعبير عـن هـذه العلاقة (الارتباط) بمصطلحات أكثر دقـة . ومـن الممكن تكوين صيغة رياضية دقيقة ، لتصف أي قيمة افتراضية لـ (y) بمعرفة أي قيمة من قيم (x) ، ويُعـد تحليل الانحـدار هو الصيغـة التقليدية لهذا النـوع من التحليل . فهـو يعيننا ـ حقيقة ـ على التنبؤ بقيمة متغير عند معرفة قيمة المتغير الآخر، ويعتمد على مجموعة المعلومات المستخرجة من بيانات مترابطة . فإذا كانت هذه البيانات غير ممثلة لمجتمع الدراسة ، أو كان الارتباط بينها مزيفاً أو غير منطقي أو بالصدفة ، فأي انحدار أو تنبؤ ينتج عنها ، يُعد غير ذي قيمة ـ أن لم يُعد مضللاً ـ مهما كانت درجة قوته .

لو استـطعنـا افـتراض أن المكتبات الفرعية الأخرى ، تطابق النمط الموصوف في عينتنا ، فقد نرى استغلال النتائج التي توصلنا إليها ، للتنبؤ بالارتباط مابين الاعارة

والنفقات في تلك المكتبات، ومثال ذلك :

من الممكن أن نتساءل: ماهي نفقات الإعارة المتوقعة فـي مكتبة ما، عندما يكون الانفاق على تكلفة الفرد تساوي ثمانية دولارات ؟، وسوف تعطينا فـي الحال المعادلة : (bx + a = y) ، نفقات تكلفة الإعارة المتوقعة، لأي من نفقات تكلفة الفـرد، فـي إطـار متـوسط قيم (x) الأصلية (3 – 7,20 دولاراً)، ولكن لـو حاولنـا التنبؤ لنفقات الإعارة، بالاستناد على تكلفة نفقات الفرد من خارج هذا المدى "Range" ، لن يمكننا التأكد من أن العلاقات نفسها (الارتباط) ستكون قائمة، حتى لو كانت هناك معقولية في افتراض أن العلاقة (الارتباط) ستكون هي نفسها خارج المدى الذي حُدد لـ (x).

فـي المثال الذي بيـن أيدينا، فـإن (y) تعتبر متغيراً تابعاً (غير مستقل)، وتُعـد أيضا القيمة غير المعروفة لنفقات تكلفة الإعارة التي نسعي إلى إيجادها، أما (x) التي تساوي 8 دولارات، فهي متغيرنا المستقل، الذي يمثل نفقات تكلفة الفرد. وشكل المعـادلة التي تستخدم في هذه الحالة كما يلي :

$$a = \frac{\Sigma y - b\Sigma x}{n}$$

وبإضافة القيم التي أصبحت ثوابت (مُحددات) في تحليل الانحدار، فالمعادلة التي تصف « انحدار Slope » الخط المنحدر، تأخذ الشكل التالي :

$$b = \frac{n\Sigma xy - (\Sigma x)(\Sigma y)}{n\Sigma x^2 - (\Sigma x)^2}$$

النتـائج :

يجب إيجاد قيمة b أولاً، حيث أن b عامل مـن العوامل المطلوبة في المعادلة الخاصة بـ a. وعليه :

حيث : $a = \frac{\Sigma y - b\Sigma x}{n}$

وحيث

$$b = \frac{n\,\Sigma xy - (\Sigma x)(\Sigma y)}{n\,\Sigma x^2 - (\Sigma x)^2}$$

$$= \frac{10\,(292,66) - (50,3)\,(52,7)}{10\,(269,61) - (50,30)^2}$$

$$= \frac{275,79}{166,01}$$

$$= 1,66$$

بالتعويض في معادلة (a):

$$a = \frac{50,30(1,66) - 52,7}{10}$$

$$= \frac{50,30(1,66) - 52,7}{10}$$

$$= \frac{-30,80}{10}$$

$$= 3,08$$

ويمكن استخدام المعادلة ($y = -3,08 + 1,66\,x$) ، في حالات التنبؤ بنفقات تكلفة الإعـارة للحـالات البحثية ، بالرغـم مـن عـدم وجـود أي بيانات عن (y) ، ولكن يمكن إيجاد القيمة التنبؤية لـ (y) ، عن طريق المعلومات التي لدينا عن قيمة (x) نفقات تكلفة الفرد .

ولكن ، قبل أن نشرع في أي إجراءات نحو تقدير قيمة (y) ، لأي حالات جديدة ، دعنا ـ أولا ـ نرى كيفية تطابق معادلة التنبؤ التي نستخدمها مع بياناتنا الأصلية .

سنستخدم أولاً قيم (x) ، لكل من العشر حالات التي لدينا للتنبؤ بقيمة (y) ، فإذا كانت قيمة (y) المتنبأة (ويرمز لها بالرمز: ($\hat{y}$) قريبة من قيمة (y) الحقيقية ، فسيدل هذا على أن المعادلة ذات فائدة جمة ، وأن بياناتنا تحتوي على علاقة خطية قوية . أما لو كانت قيم (y) الحقيقية ، مختلفة تماماً عن ($\hat{y}$) ، فإن هذا يؤكد أن النموذج الخطي أو معادلة الانحدار غير صالحة للتطبيق على البيانات ، وأن الارتباط الخطي ضعيف للغاية .

نعلم تماماً أن التنبؤ، أستناداً على البيانات التي تحت أيدينا في هذا المثال، سيكون ممتازاً للغاية، طالما أن عينة (r) قُدرت بـ 97. ومع ذلك ـ ولأغراض التدريب ـ سنقوم بعرض للعلاقة مابين (y) و (ŷ).

بالتعويض عن كل قيمة لـ (x) في المعادلة، وإتمام العملية الحسابية، سنصل إلى البيانات المدونة في جدول (14)، ولاحظ أننا أضفنا مكتبة جديدة للقائمة، أعطيت الحرف (M)، والتي تُقدر فيها نفقات تكلفة الفرد بـ 8 دولارات ولكن لا يوجد بها تقدير لنفقات تكلفة الإعارة، ونفترض غياب هذه المعلومة بسبب أن المكتبة حديثة العهد بالتشغيل. في هذه الحالة علينا افتراض وجود العلاقة الخطية خارج المدى المحدد لقيمة (x)، ولـذلـك دعنا نقول أن هذا الافتراض يبدو معقولاً، ويمكننا الآن استخدام معرفتنا بتحليل الانحدار، واستخدام أحد قيم المتغير (x) المعروفة لدينا، للتوصل إلى قمية (y) للمكتبة (M).

وبما أن البيانات الخاصة بالمتغيرين (x) و (y)، معروفة بالنسبة للمكتبات الأخرى من (A) وحتى (J)، فيمكننا استخدام هذه البيانات للحصول على قيمة (ŷ) للمكتبة (M).

حيث $\hat{y} = a + bx$

= 3,08 + 1,66 (8 دولارات)

= 10,2. كتاب للفرد. (كعينة متوقعة للإعارة، في حالة ما تكون نفقات تكلفة الفرد (x) تساوي 8 دولارات في المكتبة (M).

اتضح في نهاية السنة، أن تكلفة الإعارة للفرد في المكتبة (M)، بلغت 7,2 كتاباً وهي قيمة مختلفة تماماً عن القيمة التي تنبأنا بها، والتي بلغت 10,2 كتاباً، ويمكننا أن نرى الآن، بأن التطبيق غير المحدد (أي الأخذ بقيم من خارج المدى المُحدد لقيم (x)، يمكن أن يعطينا تنبؤات لقيم أعلى أو أقل من القيمة الحقيقية (أي تنبؤ خاطىء).

وعادة كلما كانت العلاقة قوية بين المتغيرات، كلما كان حجم الأخطاء أقل. ولو كانت العلاقة بين المتغيرات كاملة (حيث r = + 1 أو -1)، سوف لا يكون هناك أي خطأ في التنبؤ، وسيكون التنبؤ صحيحاً دائماً.

ويعـد النمـوذج الـرياضي الـذي نحاول تطبيقه على بياناتنا من النماذج الخطية، ويمكن بفحص جدول (14)، اكتشاف فرق مابين قيمة (y) الفعلية، وقيمة (ŷ) المتنبأ بها في كل مكتبة، ومن الطبيعي أننا لانود التنبؤ بالقيم الفعلية لـ (y)، لكل المكتبات

جدول (14): بيانات تتعلق بـ 11 مكتبة ، فيما يخص المتغيرات (x) و (y) وقيم ($\hat{y}$) المتنبأ بها، وحسابها عـن طريـق معـادلة الإنحـدار

الحالة	نفقات تكلفة الفرد [التغير (x)]	نفقات تكلفة الإعارة [المتغير (y)]	$\hat{y}$ قيمة (y) المتنبأ بها
A	4,60	4,5	4,556
B	4,10	4,6	3,726
C	6,70	8,2	8,042
D	3,90	2,5	3,394
E	3,00	2,1	1,900
F	5,10	4,9	5,386
G	4,00	3,9	3,560
H	7,20	8,9	8,872
I	6,20	7,7	7,212
J	5,50	5,4	6,050
M	8,00	---	10,200

من (A) إلى (J) ، إلا إذا كنا مهتمين بتحديد كمية الأخطاء في معادلة التنبؤ.

يمكن حساب عدد من الإحصاءات المختلفة، لنحدد بدقة صلاحية معادلة الانحدار لأغراض التنبؤ، بالإضافة إلى بذلك، يمكن إجراء اختبارات دلالية، على عوامل (معطيات) المعادلة، وعلى المعادلة ذاتها. لو أردت المزيد من الإيضاحات حول موضوع الانحدار، يمكن، مراجعة قائمة المراجع، في نهاية الفصل، والتي تحتوي على مراجع تعالج هذا الموضوع بتفاصيل أكثر دقة.

ومراجعة لما تم شرحه في هذا الفصل: يُعد تحليل الانحدار الخطي المتعدد، هو ذلك النوع من الانحدار المستخدم في مجال البحث العلمي، وفي هذا الصدد، قمنا بدراسة العلاقة بين متغيرات متعددة، وبالتحديد، متغيرات مستقلة متعددة (X's) ، ومتغير واحد تابع (غير مستقل (y) والهدف المعتاد من إجراءات الانحدار المتعدد أو الانحدار البسيط، هو تحديد شكل العلاقة بين المتغيرات موضع الاهتمام. ولو تمكنا من تحديد نوع العلاقة بطريقة رياضية، فيمكننا استخدام هذه المعلومات للقيام بعمل تنبؤات للحالات الفردية المقاسة على ضوء متغيرات (x) ، حيث لا يوجد بيانات تخص المتغير (y) ، الذي يمثل المتغير موضع اهتمامنا.

تمرينـــات:

1. إذا كنا نهتم بإيجاد العلاقة بين استخدام المكتبة، ومتوسط درجات امتحانات اللغة الانجليزية، حيث يبلغ معدل درجات اللغة الانجليزية بين طلبتنا 151 ، وقمنا بتشكيل، عينة مكونة من 8 طلاب، جدولنا بياناتها في الجدول المرفق. وتم قياس استخدام المكتبة عن طريق عدد الكتب التي استعارها كل طالب منهم، خلال الفصـل الـدراسي المعني (حيث تمت دراسـة منهـج اللغـة الانجليزية والاختبار فيها).

الطلاب	y عدد الكتب المستعارة	x معدل الدرجات في اللغة الانجليزية 151
A	1	70
B	4	80
C	3	70
D	4	90
E	3	80
F	6	90
G	5	70
H	3	70

2. حدد معادلة الانحدار لإيجاد العدد الذي تتنبأ به للكتب المستعارة، على ضوء معرفتك بأن معدل درجة اللغة الانجليزية يساوي 151.
3. لو كانت درجاتك تساوي 75 ، ومعدل درجات اللغة الانجليزية 151 ، فكم يكون عدد الكتب المتنبأ بها لاستعارتك؟ .

قراءات ومراجع

Blalock, Hubert M., Jr. *Social Statistics, 2d.* New York: McGraw-Hill, 1972, p. 361–383.

Ferguson, George A. *Statistical analysis in Psychology and Education.* 4th ed. New York: McGraw-Hill, 1976.

Guildford, J.P. *Fundamental Statistics in Psychology and Education.* 5th ed. New York: McGraw-Hill, 1973.

Nie, Norman, Hadlai Hull, et al. *Statistical Package for the Social Sciences.* 2d ed. New York: McGraw-Hill, 1975.

Wonnacott, Thomas H., and Ronald J. Wonnacott. *Introductory Statistics.* 2d ed. New York: Wiley, 1972.

الفصل الخامس

الاختبارات والمقاييس غير المعلمية

NON PARAMETRIC TESTS AND MEASURES

الفصل الخامس
الاختبارات والمقاييس غير المعلمية
NON PARAMETRIC TESTS AND MEASURES

تعد الاختبارات والمقاييس غير المعلمية مفيدة لأنها سهلة الاستخدام، ولا تتطلب وضع تقديرات لشكل توزيع مجتمع الدراسة للمتغير أو المتغيرات المطلوب قياسها. ولا يُشك في قدرة كثير من المتغيرات التي حُلِّلت في المكتبات ومراكز المعلومات على قياس ماهو مطلوب قياسه، ولكنها ـ رغم ذلك ـ مازالت في حاجة إلى تنقيح وتعديل قبل استخدامها في المعالجات الإحصائية الخاصة بالمكتبات ومراكز المعلومات. وسوف توضح الأمثلة التي سنعرضها في هذا الفصل، وتلك التي عرضناها في الفصول السابقة، والمتعلقة بمستويات القياس، بعض تلك المعالجات. ويمكن الاعتماد على إجراءات الاختبارات غير المعلمية عندما نكون غير متأكدين من أن مقياس الفترة (مقياس الحد) مطابق للحالة البحثية (بالرغم من أنه قد يبدو مطابقاً). ولا تتطلب إجراءات الاختبارات والمقاييس غير المعلمية افتراضات مقاييس الفترة (الحد)، ولكن يمكن تطبيقها على بيانات مقاسة بهذه المقاييس.

قد يدفعنا تحفظنا ـ في بعض الأحيان ـ إلى استخدام الإجراءات غير المعلمية، وذلك عندما يكون لدينا شك في مستوى القياس المستخدم. ومن ناحية أخرى، تمكننا هذه الإجراءات ـ غير المعلمية ـ من تجنب استخدام الإجراءات المعقدة والمفاهيم غير الواضحة، إذ يشعر بعض الناقدين المتحفظين، أن قياس المتغيرات على ضوء مفاهيم مثل: «أكثر من» أو «أقل من»، أو «عالي»، أو «متوسط»، أو «منخفض» (المتعلقة بالمقياس الاسمي)، أو مفاهيم الترتيب النظامي، مثل «الاختيار الأول»، «الاختيار الثاني» . . . «الاختيار النوني (ن)» (المتعلقة بالمقياس الترتيبي)، قد يقود إلى تطبيق معايير أقل دقة من المستوى المطلوب أو المستوى المرغوب. ومن الطبيعي أن نهتم بقياس

«الأشياء» بدقة، حتى نحصل على بيانات تقودنا إلى نتائج دقيقة وفعالة، ولكن، من ناحية أخرى، يوجد العديد من المتغيرات التي يُستعصى حلها بمقاييس الفترة (الحد)، والإجراءات غير المعلمية تفيد تماماً في هذا الأمر.

قد يكون مصطلح «غير المعلمية» مضللاً، فهو لا يعني بأن نماذجه للتحليل لا تحتوي على أي معايير أو مقاييس، ولا يعني أنها بدون فروض، بل يعني، أننا لسنا بمجبرين على وضع فروض مطولة ومشروحة عن شكل توزيع مجتمع الدراسة.

اختبارات ومعايير القياس الاسمي:

NOMINAL SCALE TESTS AND MEASURES

اختبار مربع كاي (X^2 = كا^2): The Chi-Square Test of Independante

يُعد اختبار مربع كاي (X^2) واحداً من أكثر أنواع الإختبارات فائدة، ولذلك فهو من أكثر الأدوات الإحصائية شيوعاً، ويطلق عليه في كثير من الأحيان «إحصاءات جودة التوفيق Goodness of Fit» وذلك لأنه قياس للفروقات مابين عدد مرات تكرار الظهور المشاهد (الفعلي)، وعدد مرات تكرار الظهور المتوقع (النظري) الذي تنبأنا به في فرضنا الصفري، ومعادلة X^2، تكون عادة كالآتي:

$$X^2 = \sum \frac{(fe - fo)^2}{fe}$$

حيث:

fo = التكرار المشاهد (الفعلي)

fe = التكرار المتوقع (النظري)

وهذا يعني أن مربع كاي (X^2) يساوي الفرق بين مربعات عدد التكرارات المشاهدة وعدد التكرارات المتوقعة، مقسوماً على عدد التكرارات المتوقعة، ويمكن تحويل المعادلة، لتأخذ الشكل الرياضي التالي:

$$X^2 = \sum \frac{fo^2}{fe^2} - n$$

حيث:

n = العدد الكلي للحالات أو المشاهدات.

وكما هو منطبق على كل التقنيات المُستعرضة في هذا الفصل، فمربع كاي لا يحتاج إلى مقاييس فترة (حدى)، أو حتى معايير المقايس الترتيبي، بالرغم من إمكانية تطبيقه على البيانات المُعالجة بهذه المقاييس. ويتطلب مربع كاي الآتي:-

1. عينة عشوائية.
2. كل حالة دراسية تكون مستقلة عن الحالات الدراسية الأخرى، ويعني هذا، أن لا تؤثر قيمة متغير على قيمة متغير آخر.
3. كل حالة بحثية تقع داخل خلية واحدة [الخلية، جزء محدد داخل مربع كاي].
4. يجب أن تكون 80% من عدد التكرارات المتوقعة، تساوي 5 على الأقل، أو أكبر من ذلك.

وتسمح إحصاءات مربع كاي، باختبار العلاقة بين متغيرات مجتمع ا لدراسة موضع الاهتمام، فنحن ـ هنا ـ لا نهتم بمعرفة شكل العلاقة (مثال: الارتباط الخطي)، ولكن اهتمامنا ينصب على التأكد من جودة العلاقة.

وكما كان يحدث من قبل، فنحن نفترض عدم وجود علاقة بين متغيرات مجتمع الدراسة. ويتلخص فرضنا الصفري في الآتي: المتغيران موضع البحث يعملان باستقلالية عن بعضها البعض، ونتمسك بهذا الفرض، حتى يثبت لنا العكس، عن طريق البيانات التي تمدنا بها عينة البحث.

يتطلب مربع كاي تعاملنا مع التكرار، فإذا كانت نتائجنا في شكل نسبة مئوية، فيجب تحويلها إلى تكرار، فمثلا؛ إذا كان مجموع مجتمع الدراسة يساوي 150، حيث 80% منهم نساء و 20% رجال، فمن الضروري تحويل هذه البيانات إلى 120 نساء، و 30 رجال، حتى تتناسب مع الإجراءات الحسابية لمربع كاي. ولاختبار فرضنا الصفري يجب تحديد كل من تكرارات المشاهدة، والتكرارات المتوقعة، لكل خلية (حالة بحثية)، والتكرارات المتوقعة، هي تلك التكرارات التي نتوقع ظهورها إذا لم يكن هناك أي علاقة بين متغيرات مجتمع الدراسة. وبعدها نقوم بمقارنة التكرارات المتوقعة مع التكرارات المشاهدة (الفعلية)، فإذا نتج عن المقارنة وجود فروقات كبيرة بطريقة غير عادية، نستطيع استنتاج أن افتراضنا كان خاطئاً، ونُقر بوجود علاقة بين متغيرات مجتمع الدراسة.

قد يكون من الأفضل، وقبل أن نبدأ في دراسة فعلية لمشكلة عينة ذات بيانات مُشاهدة معينة، أن نرى كيفية تقدير تكرارات متوقعة أو محتملة، وبالتالي استيعاب

منـطقية تطبيق مربـع كاي كاختبـار تحليلي للفـروقـات مابـين التكـرارات المشاهدة والتكرارات المحتملة.

لنفترض أننا انتهينا لتونا من تشكيل عينة عشوائية مكونة من 60 شخصاً بالغاً: 30 إمرأة، و 30 رجلاً، ومن بين الستين شخصاً هناك 20 يصنفون كمستفيدين من المكتبة، و 40 غير مستفيدين. وبذلك يكون لدينا متغيرين مقاسيين بعينة أفراد واحدة، هما: الجنس، واستخدام المكتبة. يمكننا وضع كل فرد في خلية واحدة فقط «في الجدولة المتقاطعة»، ويتم وضعها على الجدول بناء على القيمة التي يحصل عليها كل فرد من المتغيرين. ولنفترض أن المعلومات المعلومة لدينا تتعلق بالقيم الهامشية العامة، مثل: عدد الرجال، عدد النساء، عدد المستفيدين من المكتبة، عدد غير المستفيدين من المكتبة.

	مستفيدون	غير مستفيدين	
رجـــال	Cell خلية	Cell خلية	30
نســـاء	Cell خلية	Cell خلية	30
	20	40	60

إذا افـترضنـا عدم وجـود علاقـة بين الجنس واستخـدام المكتبـة، وأن الـرجـال لايستخدمون المكتبة أكثر من النساء، أو أن احتمال استخدام أي من الجنسين للمكتبة أكثر من الجنس الآخر، ليس وارداً. دعنا نحاول ملء خلية التوقع، بالقيم التي يمكن أن نحصل عليها إذا كان المتغيرين مستقلين عن بعضهما (أي لا يوجد بينهما علاقة).

بهذا الافتراض، وبناء على محض الصدفة وحدها، ماهي القيم التي نتوقع ظهورها في كل من الخـلايا الأربـع الفارغة؟- وكم عدد المستفيدين من الرجال؟- وعدد غير المـستفيدين منهم؟ وكم عدد المستفيدات من النســاء؟ - وكم عدد غير المستفيدات منهن؟. ونتوقع أن تكون إجابتك المنطقة كالآتي: أن نصف العدد الكلي للمجموعة

(30 من 60) من الرجال، ولذلك فأنا أتوقع أن يكون نصف عدد المستفيدين من الرجال، وعليه نقوم بحساب التكرارات المتوقعة بناء على القيم الهامشية العامة، المتمثل في مجموع الصفوف والأعمدة. وبهذا المنطق سنستنتج أن هناك 10 رجال مستفيدين، 10 نساء مستفيدات، 20 رجلاً غير مستفيد، 20 إمرأة غير مستفيدة.

يمكننا ـ لو أردنا ـ استخدام معادلة لتحديد القيم المتوقعة، بدلاً من استخدام المنطق، وفي هذه الحالة، فالقيمة المتوقعة لأي خلية تساوي مجموع الصف الذي توجد به الخلية، ويقسم الناتج على المجموع الكلي للتكرار (حجم العينة الكلي) وتأخذ المعادلة الصورة التالية:

$$\text{التكرار المتوقع} = \frac{\text{مجموع الصف الذي توجد به الخلية} \times \text{مجموع العمود الذي توجد به الخلية}}{\text{المجموع الكلي للتكرار (حجم العينة)}}$$

وعليه، فالتكرارات المتوقعة (fe) لكل خلية في المثال، تكون كالآتي:

fe للخلية a = (30) (20) ÷ 60 = 10

fe للخلية b = (30) (40) ÷ 60 = 20

fe للخلية c = (30) (20) ÷ 60 = 10

fe للخلية d = (30) (40) ÷ 60 = 20

وعلى أي حال، تثبت المشاهدة الفعلية في العينة العشوائية للمستفيدين من المكتبة، (حيث: n = 20)، أنه يوجد 30% فقط رجال، والتكرار الفعلي للرجال المستفيدين يساوي 6، ويوجد 24 غير مستفيد من الرجال، و 14 سيدة مستفيدة، 16 سيدة غير مستفيدة.

يظهر جدول (15) نتائج بيانات الدراسة المسحية، ويسجل التكرارات المشاهدة (الفعلية) في منتصف كل خلية، ويسجل أيضاً التكرارات المتوقعة (المحتملة) في الركن الأيمن الأعلى منها حتى يتمكن القارىء من إجراء المقارنات الحسابية بين المتغيرات (التكرارات).

وسؤالنا البحثي، كما يلي: هل للفرق بين التكرارات المتوقعة والتكرارات الفعلية أي دلالة إحصائية؟ـ وكما يحدث في اختبارات الفروض الأخرى، يمكننا في هذه المثال ـ اختيار مستوى دلالي، وليكن 05,. ويتم حساب مربع كاي، ببساطة وبدون أي

جدول (15): جدولة متقاطعة، لمتغيري الجنس واستخدام المكتبة، حيث، التكرار المتوقع لكل خلية مبين على الجانب الأيمن الأعلى، والتكرار المشاهد في منتصف كل خلية

	مستفيد	غير مستفيد	
رجال	10 6	20 24 6	30
سيدات	10 14 C	20 16 D	30
	20	40	60 = n

تعقيد، ويمكن تسجيل القيمة على قائمة، باتباع الآتي:

من اليمـين إلى اليسار، ونبدأ من الجانب الأيمن: الأعلى للخلية. (حيث fo = التكرار المُشاهد (الفعلي)، fe = التكرار المتوقع (النظري):

fo	fe	fo^2	fo^2 / fe
6	10	36	3,6
24	20	576	28,8
14	10	196	19,6
16	20	256	12,6
			64.8

$$X^2 = كا^2 = \sum \frac{fo^2}{fe} - n$$

$$= 64,8 - 60$$

$$= 4,8$$

يسمح لنا، الجدول الخاص بتوزيع مربع كاي، وهو نوع من التوزيعات الرياضية الخاصة، بتفسير قيم نتائج العمليات الحسابية للتحليل الإحصائي. ويوجد مثال لهذا

الجدول في «ملحق رقم (5) جدول رقم (4) » ونحتاج هذا الجدول لتحديد قيمة درجة الحرية (df) ، وكذلك لمعرفة القيمة المقابلة للمستوى الدلالي لحالتنا الدراسية (05,).

ومعادلة إيجاد درجة الحرية لمربع كاي :

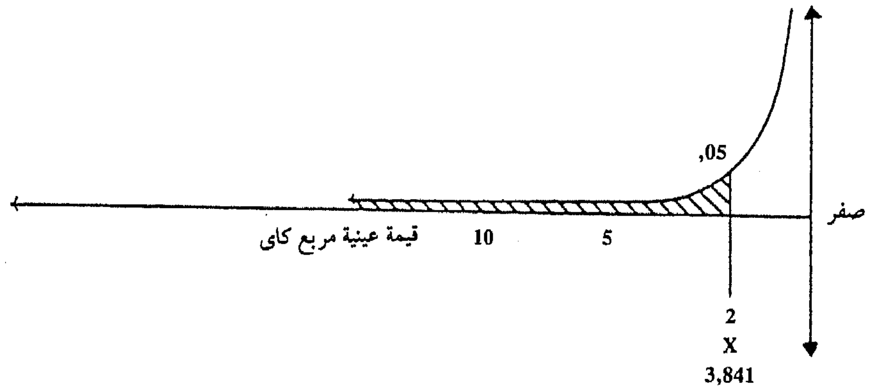
$$(1\text{–}c)\ (1\text{–}r) = df$$

= عدد الصفـوف -1 مضروباً في عدد الأعمدة ناقص واحد (حيث لكل صفين وكل عمودين درجة حرية واحدة فقط) .

وبقراءة مربع كاي (في جدول (4) بالملحق) نرى أنه في المستوى الدلالي 05, مع درجة حرية واحدة ـ إذا كان المتغير مستقلاً حقيقة ـ لا نتوقع أن تكون قيمة مربع كاي أكبر من 3,841.

شكل (21) يوضح توزيع المعاينة الإحصائية ، لإحصاءات مربع كاي ، عندما تكون متغيرات البحث حقيقة مستقلة (وهذا يعني أن الفرضية الصفرية صحيحة) ، ويمكن إيجاد قيمة هذا التوزيع ، بأخذ كل العينات المحتملة (عندما n = 60) وحساب قيمة مربـع كاي عن طريق تلك المعاينة الإحصائية ، وقد ينتج عنها احتمال ظهور بعض القيم ، واحتمال عدم ظهور البعض الآخر ، ويمكن أن نتوقع فرقاً صغيراً بين fe (التكرار المتوقع) و fo (الكرار الفعلي) ، وبالتالي نحصل على نتائج إحصائية صغيرة لمربع كاي .

تمثل المناطق المظللة (في طرف شكل (21) الظهور غير المحتمل للقيم ، والذي حددنا له قيمة تساوي 5% من مجموع المساحة الكلية التي يغطيها المنحنى . وتقع القيمة 3,841 في حافة المنطقة المظللة . وسنقوم برفض فرضنا الصفري ، لو وقعت قيمة عينتنا على



(لاحظ أن التوزيع مكون من قيم إيجابية فقط)

شكل (21): توزيع مربع كاي، بدرجة حرية واحدة 1 ، ويظهر في شكل عينة عشوائية ، مع منطقة مظللة ، بنسبة رفض 5% من المساحة الكلية

يمين القيمة 3,841 وتعد القيمة المحسوبة لـ 4,8 ، قيمة ذات دلالة إحصائية ، طالما أنها تقـع داخل المنطقة المظللة ، ويتوقع لها أن تظهر بنسبة أقل من 5% ، لوكان فرضنا صحيحاً ، ولذلك ، فنحن نرفض فرضنا الصفري ، القائل ، بأنه لاتوجد علاقة بين الجنس واستخدام المكتبة ، ونستنتج أن هذه المتغيرات مرتبطة حقيقة مع بعضها البعض .

إعـادة إلى ما سبق توضيحه بخصوص المستويات الدلالية لاختبارات الفروض ، ليس هناك شيء «مقدس» (ثابت) ، فيما يخص استخدام مستوى دلالي معين ، وبالرغم من أن المستـويات الـدلالية 05, و 01, هما أكثـر قيم المستـويات الـدلالية شيوعـاً واستخـدامـاً ، إلا أن الباحث ليس ملزماً باستخدام هذه القيم . ولذا نجد البعض يفضل استخدام قيم دلالية مختلفة عن 05, و 01, ، بغرض الخروج بمؤشرات ومعلومات أكثـر فائـدة للقـارىء . ومثـال ذلك ، احتمال الحصول على مربع كاي بقيمة 4,8 ، باستخدام درجة حرية واحدة ، يقع ما بين 05, و 01,. وفي هذه الحالة ، كنا سنرفض فرضنا الصفري إذا كان اختيارنا وقع على 05, كمستوى دلالي لحالتنا الدراسية إلا أننا كنا ـ على العكس ـ سنقبل بفرضنا الصفري لو أخذنا بـ 01, كمستوى دلالي (يقصد المؤلف بهذا المثـال ، أن اختيار قيمـة المستـوى الدلالي ؛ يتوقف على منظور الباحث لبيانات عينته ، ومدى تناسبها مع المستوى الدلالي الذي سيقع عليه الاختيار ، ومدى ملائمـة ذلـك كله للنتائج المرغوب الخروج بها من البحث ، ولذا فلسنا مجبرين على استخدام المستويات 05, و 01, المأخوذة بها في أمثلة هذا الكتاب [المترجمـان] .

قد يسـاء استخـدام اختبـار مربـع كاي في بعض الأحيان ، حيث يُعد في حكم المستحيل أن لا نحصل على «دلالة إحصائية» عندما تكون العينة كبيرة (بمعنى أننا في هذه الحالة ، نستطيع الخروج بنتائج إحصائية صحيحة ومهمة) ، فلو أمكننا ـ مثلاً ـ استخـدام قاعدة بيانات ضخمة كتلك الخاصة بإدارة إحصاءت الولايات المتحدة ، سنتمكن ـ بدون شك ـ من الحصول على كل المعلومات ـ تقريبا ـ التي تساعد على الخـروج بنتائج ذات دلالة إحصائية عن طريق استخدام مربع كاي . وبتعبير آخر ، فالمستفيد من دراستنا يجب أن يحصل على معلومات أكبر[1] .

1. يعني المؤلف بذلك : إذا لم نستطع الحصول على معلومات هامة وصحيحة باستخدام مربع كاي في حالة توافر البيانات المفصلة والدقيقة ، فإننا نكون قد أسأنا استخدام مربع كاي . (المترجمان) .

مثال لذلك، عينة مكونة من 50 شخصاً بالغاً الغرض منها إظهار مدى استخدام المستفيدين من الجنسين للمكتبة، وتم التعبير عنها في الجدولة المتقاطعة التالية:

	مستفيد	غير مستفيد	
ذكور	12,5 10	12,5 15	25
أناث	12,5 15	12,5 10	25
	25	25	50

fo	fe	fo^2	$\frac{fo^2}{fe}$
10	12,5	100	8
15	12,5	225	18
15	12,5	225	18
10	12,5	100	8
			52

$$X^2 = 52 - 50 = 2$$

نتيجة إحصائية غير دلالية، في المستوى الدلالي 05, ودرجة حرية (df) تساوي 1، الفرض الصفري لم يرفض، حيث لا يوجد لدينا ما يجعلنا نفترض أن هناك فرقاً في استخدام المكتبة بين الجنسين.

نلاحظ في هذا المثال، أن 40% من الرجال في عينتنا هم مستفيدون من المكتبة ($\frac{10}{25}$ = ,40)، و60% منهم غير مستفيدين ($\frac{15}{25}$ = ,60)، 60% من النساء مستفيدات من المكتبة ($\frac{15}{25}$ = ,60)، و40% من النساء غير مستفيدات ($\frac{15}{25}$ = ,40). وهذه النتائج قد تدفعنا لإجراء اختبار على الفرض القائل بأن للجنس علاقة باستخدام المكتبة. ولنعتبر الآن، أننا أعدنا الاختبار، بعد أن قمنا بزيادة بحجم العينة لتصبح 5000 شخص بالغ.

	مستفيد	غير مستفيد	
ذكور	1,250 1000	1,250 1500	2500
اناث	1,250 1500	1,250 1000	2500
	2500	2500	

fo	fe	fo^2	$\frac{fo^2}{fe}$
1000	1,250	1,000,000	800
1500	1,250	2,250,000	1,800
1500	1,250	2,250,000	1,800
1000	1,250	1,000,000	800
			5,200

$$X^2 = 5200 - 500 = 200$$

الدلالة الإحصائية في المستوى الدلالي 01, تفرض علينا رفض فرضيتنا الصفرية، واستنتاج أن هناك علاقة بين الجنس واستخدام المكتبة.

من الواضح أن الشيء الوحيد الذي تغير، هو حجم العينة، أما نسب استخدام الجنسين للمكتبة فلم تتغير.

وعلى أي حال، يجب أن نتذكر، أن تلخيص البيانات عن طريق النسب المئوية، أو التناسب، يساعد إلى حد كبير، وأن التزام التحفظ الشديد ـ أمر مرغوب فيه ـ عند استخدام مربع كاي مع العينات الكبيرة.

لا يقتصر استخدام مربع كاي، على الجدولة بصفين وعمودين فقط، (كالجداول السابقة)، بل قد يمتد استخدامه ليشمل جداول بيانات أكبر حجماً لتشتمل على أكثر من متغيرين، ومن حوالي سنتين قام موريس مارشانت Maurice Marchant ، بتحليل

استخدام أعضاء هيئة التدريس للمكتبات العامة، في آن آربر Ann Arbor، بولاية ميتشيجان Michigan[2]، وقد أشارت نتائجه، أن 42 من 113 عالما نفسياً، و 17 أخصائياً من 68 أخصائياً في علوم الحياة، و 33 مهندساً من 203 مهندس، و 20 أستاذا من 78 أستاذاً للغة الانجليزية، يستفيدون من المكتبة العامة. وكانت النسب المئوية كالآتي: 37% في علم النفس، و 25% في علم الأحياء، و 16% في الهندسة، و 26% في اللغة الانجليزية. والتساؤل هنا: هل تعد هذه النتائج ذات دلالة إحصائية ؟- يؤكد مارشانت على أنها كذلك. ونستطيع أن نتأكد من إدعائه، بإخضاع بياناته، لجدولة إحصائية، نطبق فيها اختبار مربع كاي، وسيكون شكل الجدول كالآتي:

	علم النفس	علم الأحياء	الهندسة	اللغة الانجليزية	
مستفيدون	27 / 42	16 / 17	49 / 33	19 / 20	112
غير مستفيدون	86 / 71	52 / 51	154 / 170	69 / 58	350
	113	68	203	78	462

الرقم المسجل في مركز كل خلية، يمثل عدد المستفيدين أو عدد غير المستفيدين من المكتبة في كل تخصص، أما الرقم المسجل في الجانب الأيسر الأعلى، فيمثل عدد الأشخاص الذي نتوقع استفادتهم من المكتبة عن طريق الصدفة، وسنقوم بحساب القيم المتوقعة (fe) لغير المستفيدين في مجال علم النفس واللغة الانجليزية، بغرض تذكر العمليات الحسابية التي تجري في هذا الصدد:

$$fe = \frac{(113)(350)}{462} = 85,6$$ غير مستفيد في مجال علم النفس

$$fe = \frac{(78)(350)}{462} = 59,0$$ غير مستفيد في مجال اللغة الانجليزية

2. Maurice MARCHANT "Faculty as Public Library Patrons" Wilson Library Bulletin, January 1969, P. 445-446.

وفيما يلي، إحصاءات مربع كاي، للعينة العشوائية البحثية :

$\frac{fo^2}{fe}$	fo^2	fe	fo
65,3	1764	27	42
18,1	289	16	17
22,2	1089	49	33
21,1	400	19	20
58,6	5041	86	71
50,0	2601	52	51
187,7	28900	154	170
57,0	3364	59	58
480,0			462

مربع كاى $[\ 18 = 462 - 480{,}0 = X^2\ ]$

قوة الارتبــــاط Strength of Association

هنـاك الكثـير من المقـاييس التي يمكن أن تساعدنا على معرفة قوة الارتباط بين متغيرين، طالما أنهما (أي المتغيرين)، مرتبطين فعليا وبقوة. ترتبط كل المقاييس التي تناقش هنا، بمربع كاي، ولذلك فهي محدودة. ويُعد مربع في Phi-Square ، ويرمز له بالرمز ($\varnothing^2$) (من أبسط المقاييس المستخدمة، ويتطلب إيجاده، مجرد قسمة مربع كاي على قيمة n ($\varnothing^2 = \frac{X^2}{n}$) في كل الأمثلة السابقة التي وردت عاليه، عندما ظلت قيمة النسب كما هي بدون تغير، وبينما زادت حجم العينة، فإن قيمة ($\varnothing^2$) ثابتة لم تتغير، ولذلك فإن مقياس $\varnothing^2$) يساعد في الرقابة والسيطرة على حجم العينة المؤثرة على الدلالة الإحصـائية للنتـائـج . في جدول 2x2 (أي المكـون من عمـودين وصفين) قد تصل قيمة $\varnothing^2$ إلى (1) صحيح ، والجدولة المتقاطعة التالية ، تظهر لنا علاقة مثالية لـ $\varnothing^2$.

12,5 صفر	12,5 25	25
12,5 25	12,5 صفر	25
25	25	50

تقدر قيمة مربع كاي بـ 50. وعليه تقدر قيمة $\varnothing^2 = \frac{X^2}{n} = \frac{50}{50} = 1$ صحيح

(علاقة مثالية) هناك مقاييس أخرى تتضمن:

$$T^2 = \frac{\varnothing^2}{\sqrt{(\text{ع}-1)(\text{ر}-1)}}$$

و

$$V^2 = \frac{\varnothing^2}{\sqrt{(\text{ع}-1)(\text{ر}-1)}}$$

(استخدام عدد الصفوف، وعدد الأعمدة أيهما أقل)

في جدولة 2x2 (أي مكونة من صفين وعددين)، قيم $\varnothing^2$ ، T^2 ، V^2 تكون مساوية لبعضها البعض.

معامل التوفيـــق The Contingency Confficient C[3]

يُعرف ـ أيضا ـ باسم «بيرسون (C) ، Pearson's C » أو معامل توفيق بيرسون Pearson's Contingency Confficient ، ويعتمد ـ أيضاً ـ على مربع كاي، ويُعد من أكثر اختبارات المجموعة استخداماً وانتشاراً. وفيه، تقدر قيمة (C) بصفر، عندما لا يكون هناك أي ارتباط بين المتغيرات، أوعندما تكون المتغيرات مرتبطة ارتباطا تاماً (كاملاً). أو تعتمد على بعضها البعض. ولا يكون لـ (C) أي وحدة كقيمة قصوى.

3. تشير إليه بعض المراجع بـ: ك: «معامل حُسن المطابقة». (المترجمـان).

في حالات الجدولة (2x2) ، القيمة القصوى لـ C = 7.7 ، في الجدولة (3x3) تكون قيمتها = ,816 وحقيقة الأمر، عندما تتساوي أعداد الصفوف والأعمدة فإن الحد الأقصى لـ (C) يكون:

$$C = \sqrt{\frac{\text{عدد الأعمدة} - 1}{\text{عدد الأعمدة}}}$$

يؤخذ بمقياس (C) - كما يحدث مع كثير من المقاييس - لأن تطبيقه يتطلب الحد الأدنى من المعايير، ويسهل حسابه. ويمكن مقارنة معاملات التوفيق مع بعضها البعض، عندما يعتمدون جميعاً على جداول من نفس الحجم (أي، جداول تحتوي على نفس العدد من الصفوف والأعمدة) ولكن لا يمكن مقارنتها بصورة مباشرة، مع مقاييس الارتباط الأخرى المخصصة للتعامل مع القياسات الإسمية أو الحدية، مثل معامل ارتباط بيرسون (r) ، ومعامل ارتباط سبيرمان rs).

ومعادلة معامل التوفيق تكون كالآتي:

$$C = \sqrt{\frac{X^2}{n + X^2}}$$

نتج من تقدير نتائج متغيرين، أن قيمة مربع كاي = 15 ، n = 50 ، مما ينتج عنه أن قيمة C = ,48 . فإذا ماكانت القيمة القصوى المحتملة لـ (C) في جدولة (2x2) (أي جدول متغيرين)، تساوي ,707. فإن قيمة (C) في هذه الحالة (,48) تُعد قيمة ارتباط عالية. ولمثالنا السابق عن مربع كاي، عندما وجدنا علاقة دلالية بين الجنس واستخدام المكتبة، يمكن تطبيق معامل التوفيق عليها، وستكون قيمة C = ,20 ، وهي قيمة متواضعة للغاية، إذا علمنا أن قيمة (C) في هذه الحالة يمكن أن تصل إلى ,707.

مقاييس واختبارات القياس الترتيبي

ORDINAL SCALE TESTS AND MEASURES

اختبار وولد ـ ولفوتز (أو اختبار Run)[4] The Wald - Wolfowitz Runs Test

يمكن استخدام هذا الإجراء، عندما نريد اختبار فرض صفري، في حالة وجود عينتين مأخوذتين من نفس المجتمع الدراسي. ويتطلب ذلك، أن تكون لدينا عينات عشوائية مستقلة، وعلى الأقل معيار للقياس الترتيبي. ولكن لا نهتم ـ هنا ـ إذا ما كان مجتمع أو مجتمعات الدراسة موزعة توزيعاً طبيعياً أو غير طبيعي.

يُعد اختبار الـ Runs ـ في الواقع ـ من الاختبارات الشكلية المتعاملة مع مجتمعات دراسية عالية التشتت، ويُعد ـ أيضاً ـ من اختبارات النزعة المركزية Central Tendency، واختبارات التغاير. ولكنه في نفس الوقت لا يحدد الفروقات في مقياس واحد، كما يحدث عادة ـ مع اختبارات النزعة المركزية والفرض الصفري هنا، يعني أن العينتين مستخدمتين من مجتمعات دراسية متساوية التوزيع، ونفترض أن هذه حقيقة، حتى يثبت لنا العكس عن طريق تحليل بيانات العينة.

دعنا نفترض أننا في حالة بحثية لعينة، قمنا بسؤال الطلاب بترتيب عينة عشوائية مكونة من 16 كتاباً مرجعياً بناء على أهميتها بالنسبة للطلاب. ويراعى في ترتيب الكتب، أن ترتب على القائمة حسب أهميتها (حيث: يكون الكتاب رقم (1) هو الأكثر أهمية، وكتاب رقم 16 هو أقلها أهمية)، وتتألف الستة عشر كتاباً التي سيقوم الطلاب بترتيبها من 8 كتب مختارة عن طريق أعضاء هيئة التدريس، والـ 8 الأخرى، أُختيرت عن طريق المسؤولين عن المكتبة، وسؤالنا هو هل هناك فروقات ذات دلالة وأهمية، بين الكتب المرجعية التي اختارها أعضاء هية التدريس، وتلك أُختيرت عن طريق المكتبة؟.

4. كلمة Run هنا تعني ظهور حدث أو مجموعة من الأحداث المتشابهة التي تسبقها أو تعقبها أحداث تختلف عنها، مثال لذلك: إذا كانت النتائج التالية ناتجة عن إلقاء عملة معدنية عشر مرات بحيث (ظ) ترمز إلى ظهور العملة و (و) ترمز إلى وجه العملة:

و و	ظ	و	ظ ظ ظ	و	ظ	و
1	2	3	4	5	6	7

نلاحظ أن النتائج تبدأ بظهور وجهين «و، و»، ثم ظهر «ظ»، «و» . . . الخ. يطلق على كل مجموعة من الوجوه أو الظهور المتشابهة، والتي تسبقها أحداث تختلف عنها اسم Run وواضح من المثال أن لدينا Runs 7 (يمكننا القول بأن: Run تعني ظهور الحدث بين فواصل من أحداث مختلفة). (المترجمان).

وافترضنا الصفري ينص على أنه لا توجد فروقات ذات دلالة إحصائية بين اختيارات أعضاء هيئة التدريس واختيارات المكتبيين ؟ .

وبإجراء اختبار يتطلب الحد الأدنى من العمليات الحسابية . فيما لو كان مجموع $n_1 + n_2$ يساوي أو يقل عن 20 ، (استخدم جدول (5) في الملحق)، وفي حالة أن مجموع n_1 n أكثر من 20 ، فتوزيع المعاينة العشوائية لـ R ، يكون ـ تقريبا ـ طبيعيا، ويمكن استخدام المعادلة التالية :

$$\text{المتوسط الحسابي} = \mu_R = \frac{2n_1n_2}{n_1+n_2} + 1$$

$$\text{الانحراف المعياري} = \sigma_R = \sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}$$

في حالتنا البحثية، مجموع n = 16 ، وهذا أقل من 20 ، ولذا يمكننا استخدام جدول القيم الحرجة لـ (R) ، في اختبار الـ Runs بمستوى دلالي 05,. ولنفعل ذلك، يجب أولا، ترتيب الكتب حسب أهميتها، مع وضع خطوط أسفل الكتب التي أختيرت عن طريق المكتبيين، ووضع خطوط أعلى الكتب التي أُختيرت عن طريق أعضاء هيئة التدريس، دعنا نفترض أننا حصلنا على النتائج التالية :

$$\overline{16\ \ 15\ \ 14\ \ 13}\ \ \underline{12\ \ 11}\ \ \overline{10}\ \ \underline{9}\ \ \overline{8}\ \ \underline{7\ \ 6\ \ 5\ \ 4}\ \ \overline{3\ \ 2}\ \ \underline{1}$$

حتى هذه المرحلة لم نفعل سوى وضع خطوط لتمييز الكتب التي أُختيرت عن طريق أعضاء هيئة التدريس أو عن طريق المكتبيين. ويتضح ، أن الكتاب الذي يُعد أكثر أهمية ، ثم اختياره عن طريق المكتبيين ، الكتاب الثاني والثالث من ضمن المجموعة التي أُختيرت عن طريق أعضاء هيئة التدريس . . . وهكذا. ونحصل على عدد الـ Runs ، عن طريق حساب عدد الخطوط ، كل خط واحد يساوي Run واحد. وسنجد لدينا 8 Runs. وباستخدام جدول (6) في الملحق، نراجع n عند الرقم 8 حيث حجم عينتنا n = 8) ، وعبر الصف n نذهب إلى الرقم 8 (حيث حجم عينتنا n = 8) ، ونجد أنه باستخدام عينات من هذا الحجم يجب أن نتوقع عدداً من الـ Runs يساوي 5 أو أقل، حتى يكون المستوى الدلالي مساوياً لـ 05,. وحيث أن نتائج اختبارنا تشير إلى أن R = 8 Runs ، وهذا أكبر من القيمة المتنبأ بها في توزيع الجدول، نستنتج أن R = 8 ليست

ذات دلالة إحصائية. ولذلك نحن لا نرفض فرضنا الصفري، عندما يكون المستوى الدلالي مساويات 05,.

تُعد معطيات المثال السابق، من الصغر بحيث أنها لاتحتاج إلى تطبيق المعادلات التـالية، (التي يفترض استخدامها عندما تتجاوز قيمة ($n + {}_2n$:20)، ولكن يمكن استخدام المعادلة التالية، بغرض الشرح والإيضاح:

$$\text{المتوسط الحسابي} = \mu_R = \frac{(8X8)2}{1-8} + 1 = \frac{128}{16} + 1 = 9$$

$$\text{الانحراف المعياري} = \sigma_R = \sqrt{\frac{[8-8-(8x8)2](8x8)2}{(1-8+8)^2(8+8)}} = 1,93$$

$$\text{الدرجة المعيارية} = Z = \frac{\mu_R - R}{\sigma_R} = \frac{9-8}{1,93} = \frac{1-}{1,93} = ,5- \quad (\text{تقريبا})$$

باستخدام جدول المناطق الواقعة أسفل المنحنى الطبيعي، نجد أن Z عند المستوى الدلالي 05, أكبر من 1,96 وفي هذه الحالة، لا نرفض فرضنا الصفري (نرفض فرضنا الصفري في حالة ما إذا كانت (Z) تكون مساوية أو أصغر من القيمة 1,96).

ودعنا نتخيل، كيف يكون شكل الـ Runs، عندما تكون هناك دلالة إحصائية، سنحصل على نتائج شبيهة بالتالي:

16 15 14 12 11 10 9 8 7 6 5 4 3 2 1

لو حدث هذا، لكنـا قد حصلنا على عدد 2، Runs (بمعنى أن كل اختيارات الكتب التي قام بها المكتبيون تفوقت في الترتيب على الاختيارات التي قام بها أعضاء هيئة التدريس). وهذا يعني، كلما زاد عدد الـ Runs كلما زاد احتمال عدم الارتباط العشوائي بين العينتين.

معامل ارتباط سبيرمان (r s) **Spearman's Rank-Order Correlation (r s)[5]**

يسمح لنا هذا الإجراء بقياس قوة الارتباط بين متغيرين مقاسين ترتيبياً، وعندما يكون الترتيب متوافقاً تماما، فإن القيمة القصوى لـ rs = + 1 ، وعندما يكون الترتيب مختلفاً تماما، تكون القيمة القصوى لـ rs = - 1.

كمثال لذلك، لنفترض أننا سألنا الطلاب المتخرجين وغير المتخرجين، القيام بترتيب أولياتهم في كيفية قضائهم لأوقات فراغهم، وحددنا الإختيارات بالآتي: الكتب، التلفزيون، الراديو، المجلات، الأفلام السينمائية. ومن خلال إيجاد قيمة r S ، سنحاول معرفة مدى تقارب وجهات نظر الطلاب المتخرجين وغير المتخرجين، حول قضاء وقت فراغهم، ولنفترض أن نتائجنا كانت على النحو التالي:

الأنشطة	غير متخرجين	المتخرجين	D_i	D_i^2
تليفزيون	1	2	-1	1
أفلام سينمائية	2	3	-1	1
كتب	3	1	2	4
مجلات	4	4	0	0
راديو	5	5	0	0
حيث D =الفرق				6

اختار غير المتخرجون التلفزيون كأفضلية أولى لقضاء وقت فراغهم، بينما اختار المتخرجون التلفزيون كأفضلية ثانية . . . وهكذا. ويشير عمود Di إلى الفرق بين ترتيب الأفضليات للمجموعتين، وتحسب قوة الارتباط، وقوة التوافق بين الترتيبات بالمعادلة التالية:

$$rs = 1 - \frac{6\Sigma(D_i)^2}{n(n^2-1)} = 1 - \frac{6(6)}{5(25-1)} = .70$$

حيث n = المجموع الكلي للمتغيرات المرتبة، ولاغرابة في حصولنا على هذا الارتباط المتواضع، فكما نرى من فحص بيانات هذا الجدول، أن المتخرجين وغير المتخرجين

5. يطلق عليه أيضاً: معامل الارتباط الترتيبي لسبيرمان. (المترجمان).

اتفقوا في الرأي على شيئين - فقط - من حيث الأفضلية في قضاء وقت الفراغ (المجلات، الراديو).

وللمزيد من إلقاء الضوء على كيفية عمل r s وتفسيرها، دعنا نرى ماذا كان سيحدث في حالة «عدم التوافق المطلق»، وأن r s = - 1.

الأنشطة	متخرجين	اللامتخرجين	Di	Di^2
نشاط A	1	5	4	16
نشاط B	2	4	2 -	4
نشاط C	3	3	0	0
نشاط D	4	2	2 -	4
نشاط E	5	1	4	16
				40

$$r_s = 1 - \frac{6(40)}{5(25-1)} = 1 - \frac{240}{120} = 1 - 2 = -1$$

في هذه الحالة قام المتخرجون وغير المتخرجون بترتيب قائمة بكامل الأفضليات، بطريقة معاكسة تماما. وتشير r s إلى الفرق من خلال قيمتها التي تساوي -1 ، وإن كان هذا لا يعطي النسبة المئوية للتوافق بين المجموعين إلا أنه يدلنا على درجة الترابط بين ترتيب مجموعتي الأنشطة، وفي هذه الحالة، هناك «عدم توافق كلي» بين المجموعتين، بالرغم من وجود اتفاق على نوع واحد من أنواع الأنشطة (نشاط C).

استخدام r s مع بيانات القياس الحدي: Use of r s with Interval Scale Data

تكون لدينا أسبابنا في بعض الأحيان - للشك في «دقة المقاييس» المستخدمة في اختبارات القياس الحدى، ومثال لذلك: الحجم الفعلي لمقتنيات مكتبة كبيرة دائمة مشكوك في تقديره. وبتقديرنا للصعوبات التي تواجه الباحث للحصول على تقدير دقيق لمثل هذه المجموعات، يمكن أن نتساءل عن ماهية الفرق بين مقتنيات مكتبة تقدر بـ 2332650 مجلد، وأخرى تقدر مقتنياتها بـ 2340965 مجلد. بالإضافة، إذا ما ربطنا بين حجم المقتنيات ومتغير آخر كبير الحجم، كالميزانية التأسيسية - مثلا - سيصبح إجراء العمليات الحسابية في غاية التعقيد. واستخدام r s في هذه الحالة قد يكون ذو فائدة

كبيرة، ولتوضيح هذا الأمر، نستطيع قياس الارتباط مابين حجم مجتمع المستفيدين من المكتبة وبين حجم مقتنيات المكتبة، ولنا أن نتساءل: هل يتزايد حجم مقتنيات المكتبة، بازدياد حجم مجتمع المستفيدين؟. والبيانات التالية تخص عدد 10 مكتبات تتبع النظام المحلي.

المكتبة	الحجم الكلي لمجتمع المستفيدين	الترتيب	الحجم الكلي لمقتنيات المكتبة	الترتيب	Di	Di
A	38,731	10	55,756	10	0	0
B	66,281	8	83,866	7	1	1
C	125,916	2	109,527	5	– 3	9
D	103,624	5	106,840	6	– 1	1
E	46,228	9	83,404	8	1	1
F	181,097	1	282,081	1	0	0
G	102,676	6	79,861	9	– 3	9
H	79,950	7	147,568	3	4	16
I	107,930	4	151,176	2	2	4
J	121,692	3	114,098	4	– 1	1
						42

$$r_s = 1 - \frac{6(\sum D_i^2)}{n(n^2-1)} = 1 - \frac{6(42)}{10(99)} = 1 - \frac{252}{990} = 1 - ,25 = ,75$$

ويعتبر إجراء العمليات الحسابية، لقياس معامل بيرسون r s بدون الاستعانة بالوسائط الآلية (كالحواسيب)، أمراً في غاية الصعوبة، حيث يتطلب تعاملنا مع أعداد قد تصل إلى 12 رقماً. وتبرز النتائج (التي حصلنا عليها عن طريق حسابها بالحاسب الآلي)، أن قيم r = ,80، ولا يُعد هذا أكبر بكثير من قيمة r s التي تم الحصول عليها بسهولة (بدون استخدام الإمكانات الآلية). ولذا، فعندما لا تتوفر لنا الإمكانات الآلية، أو الوقت الكافي، للتعامل مع العمليات الحسابية لا ستخراج قيمة r، يمكننا الاكتفاء بقيمة rs التي نحصل عليها بسهولة أكثر، وبدون الحاجة إلى الإمكانات الآلية.

اختبار تو لكيندال Kendall's Tau

يستخدم هذا الإجراء أيضا، للحصول على مقياس الارتباط بين المقاييس الترتيبية، ويتطلب أن نحصل على الأقل، على مقياس ترتيبي واحد. واختبار Tau يشبه إلى حد كبير اختبار r s ، حيث يتراوح مدى قيمته من + 1 إلى − 1. وبتطبيقه على نفس المثال السابق المستخدم لـ R s يمكن بذلك المقارنة بين المقياسين.

e	d	c	b	a	
راديو	مجلات	كتب	أفلام	تلفزيون	
5	4	3	2	1	غير متخرجون
5	4	1	3	2	متخرجون

وسؤالنا هو: ماهي الأنشطة الثنائية المتشابهة، التي رتبها المتخرجون وغير المتخرجين بنفس الاتجاه أو النظام المتبع في الجدول؟

تم ترميز الأنشطة على النحو التالي: a ، b ، c ، d ، e. ولتسهيل فحص الأنشطة ثنائياً، ولتحديد قيمة Tau ، يجب علينا ترتيب نظام مجموعة واحدة على الأقل (ولتكن غير متخرجين) من الأدنى إلى الأعلى، (وليس من الضروري أن نقوم بهذا العمل عند حسابنا لـ rs) ، ونعطي كل ثنائي متفق قيمة تساوي = 1 ، وكل ثنائي غير متفق قيمة تساوي −1.

مثال الثنائي a ، b (التلفزيون، والأفلام)، مرتبان بنفس النظام، بمعنى أن المتخرجين وغير المتخرجين رتبوا التلفزيون في رتبة تعلو رتبة الأفلام، وكل الاحتمالات الثنائية في هذا المثال، قيست للحصول على الإحصائية S ، كما هو موضح في التالي:

2+=1+1+1-1+ = e ، a + d ، a + c ، a + b ، a

1+= 1+1+1- = e ، b + d ، b + c ، b

2+= 1+1+= e ، c + d ، c

1+= 1+= e ،d

----

6 = S

$$,60 = \frac{6}{10} = \frac{6}{\frac{20}{2}} = \frac{6}{\frac{(4)(5)}{2}} = \frac{S}{\frac{(1-n)n}{2}} = tau$$

يبدو أن قيمة Tau ,60 تشير إلى قوة ترابط منخفضة، أكثر من قيمة r s ، التي كانت تساوي 70, وهذا ما يحدث دائماً عندما نربع الفروقات بين الرتب، حيث يميل r s إلى إعطاء قيمة أكبر نسبياً للفروقات القصوى، بينما تعطي Tau ، قيمة متكافئة لكل الفروقات.

ودعنا نعطي مثالاً لاختبار Tau ، في أحد البرامج التدريبية حول الإدارة بالأهداف، تم ترتيب الأهداف التالية عن طريق المشتركين في البرنامج من الكوادر الفنية بالمكتبة والممثلين للإدارة العليا بها: مرتبات أكبر، تخفيض العمالة، نسبة إنتاج أكبر، مسؤولية ومزيد من الوضوح في خطوط السلطة، تحسين فوائد الخدمة الإضافية، إدارة قوية ومعتدلة.

ترتيب الأهداف

	a إنتاج أعلى	b إدارة قوية ومعتدلة	c مسئولية وسلطة	d عمالة أقل	e فوائد إضافية	f مرتبات
الإدارة	1	2	3	4	5	6
العاملين	5	4	2	6	3	1

a ، b + a ، c + a ، d + a ، e + a ، f = −1 −1 +1 −1 −1 = −3

b ، c + b ، d + b ، e + b ، f = −1 +1 −1 −1 = −2

c ، d + c ، e + c ، f = +1 +1 −1 = +1

d ، e + d ، f = −1 −1 = −2

e ، f = −1 = −1

S = −7

$$tau = \frac{S}{\frac{n(n-1)}{2}} = \frac{-7}{\frac{(6)(5)}{2}} = \frac{-7}{\frac{30}{2}} = \frac{-7}{15} = .47$$

وتشـير حسـابـاتنا لقيمة Tau ، بوجود ترابط عكسي أو سلبي: فقد قام كل من الإداريين، والعاملين بالمكتبة، بترتيب الأحداث بطريقة مخالفة عن طريقة الآخر، وحسـاب Tau يكون صعباً وغير عملي (بالطريقة التي عُرفت هنا) عندما يكون عدد الأشياء المطلوب ترتيبها كبيراً.

(أنظر مرجع جلاسي وستانلي Glass and Stanley في نهاية هذا الفصل).

لقـد تجنبنـا منـاقشة مشكلة مهمة في هذا الفصل، والخاصة بالترتيب المتقارب، وبعبـارة أخـرى ماذا نفعـل إذا ماكان هناك شيئين مطلوب ترتيبها، ثم أعطيا نفس الترتيب؟. في مثالنا السابق، ماذا كان سيحدث، لو لم يتمكن الإداريون من التمييز بين مسؤوليتهم تجاه العمل ومسؤوليتهم تجاه العاملين، في ترتيبهم للأحداث، وشعروا بأنه لا فرق هناك من زاوية الأهمية ـ بين مصالح العمل ومصالح العاملين؟.

عنـدما يظهر مثل هذا الترتيب المتقارب، في حساب قيمة Tau أو قيمة r s تتناثر المعاملات، وتطرأ الضرورة لإجراء بعض التعديلات لتصحيح هذا الترتيب المتقارب.

لو كنا ملزمين بالخروج بمؤشرات استنتاجية من عينة إحصائية لمعلمه مجتمع دراسي، عن طريق اختبار r s أو Tau ، فنحن مجبرون على استخدام إجراءات اختبار الفروض، وقد تخيرنا عدم التعرض لهذه المواضيع العالية التخصص. في عملنا هذا، ولجأنا بدلاً من ذلك إلى إحالة القارىء لمراجع متخصصة، قمنا بتدوينها في نهاية الفصل.

(أنظر هايز و سيجل **Hays and Siegel**)

تستوجب الإجراءات التي تعرضنا لها في هذا الفصل لبعض المتطلبات من جانب الباحث، أولا، يجب علينا ألا نشغل أنفسنا بصلاحية المناهج البحثية، إذا لم نكن على علم كافي بشكل التوزيع لمتغيرات مجتمع الدراسة، ولم يكن لدينا على الأقل مستوى حدى واحد من المعايير، ثانياً، من الطبيعي، إذا كانت مقاييس وإجراءات المعلمة صالحة للتطبيق، يجب علينا ـ في هذه الحالة ـ محاولة تطبيقها، فالمناهج المعلمية أقوى بكثير في نتائجها من المناهج غير المعلمية.

وبعبارة أخرى، إذا كان هناك فرق واضح بين مجموعاتنا البحثية، أو إذا كان هناك

علاقة بين متغيرات بحثنا، فمن الأفضل أن نتعامل معها ونحللها عن طريق إجراءات الاختبارات المعلمية.

وعلى أية حال، في كثير من الحالات، فإننا لا نشعر بأن استخدام الإجراءت المعلمية سيكون مفيداً، وقد يرجع ذلك إلى نقص في المعلومات أو البيانات، أو إلى ضعف في مستوى أدوات القياس. وفي مثل هذه الحالات، فإن استخدام الإجراءات غير المعلمية يكون مفيداً.

أننا نشعر بأن هذه التقنيات في غاية الفائدة للعاملين في مجال المكتبات، فهي عبارة عن عينة صغيرة من التقنيات المصنفة كإحصاء غير معلمي. ونأمل بأن نكون قد أعطينا القاريء، المعلومات الأساسية، حول كيفية الاستخدام والإفادة من هذه الإجراءات، ونحث القراء على استخدامها في حياتهم المهنية، للتعرف عليها وعلى مدى ملاءمتها لاحتياجاتهم بطريقة أفضل.

المراجع والقراءات:

Babbie, Earl R. *The Practice of Social Research*. Belmont, Calif.: Wads-worth, 1975.

Blalock, Hubert M., Jr. *Social Statistics*. 2d ed. New York: McGraw-Hill, 1972, p. 249-254, 275-287, 291-302, 418-426.

G lass, Gene, and Julian C. Stanley. *Statistical Methods in Education and Psychology*. Englewood Cliffs, N.J.: Prentice-Hall, 1970.

Hays, W. L. *Statistics for the Social Sciences*, 2d ed. New York: Holt, Rinehart and Winston, 1973.

Kerlinger, Fred N. *Foundations of Behavioral Research*. 2d ed. New York: Holt, Rinehart and Winston, 1973.

Siegel, Sidney. *Non-Parametric Statistics*. New York: McGraw-Hill, 1956.

ملحق الجداول

1 - جدول الأرقام العشوائية .

2 - جدول المساحات تحت المنحنى الطبيعي .

3 - جدول توزيع t.

4 - جدول توزيع مربع كاي X^2

5 - جدول معامل الارتباط .

6 - جدول القيم الحرجة لـ R في اختبار Runs.

جدول (1): الأرقــام العشوائية

94015	46874	32444	48277	59820	96163	64654	25843	41145	42820
74108	88222	88570	74015	25704	91035	01755	14750	48968	38603
62880	87873	95160	59221	22304	90314	72877	17334	39283	04149
11748	12102	80580	41867	17710	59621	06554	07850	73950	79552
17944	05600	60478	03343	25852	58905	57216	39618	49856	99326
66067	42792	95043	52680	46780	56487	09971	59481	37006	22186
54244	91030	45547	70818	59849	96169	61459	21647	87417	17198
30945	57589	31732	57260	47670	07654	46376	25366	94746	49580
69170	37403	86995	90307	94304	71803	26825	05511	12459	91314
08345	88975	35841	85771	08105	59987	87112	21476	14713	71181
27767	43584	85301	88977	29490	69714	73035	41207	74699	09310
13025	14338	54066	15243	47724	66733	47431	43905	31048	56699
80217	36292	98525	24335	24432	24896	43277	58874	11466	16082
10875	62004	90391	61105	57411	06368	53856	30743	08670	84741
54127	57326	26629	19087	24472	88779	30540	27886	61732	75454
60311	42824	37301	42678	45990	43242	17374	52003	70707	70214
49739	71484	92003	98086	76668	73209	59202	11973	02902	33250
78626	51594	16453	94614	39014	97066	83012	09832	25571	77628
66692	13986	99837	00582	81232	44987	09504	96412	90193	79568
44071	28091	07362	97703	76447	42537	98524	97831	65704	09514
41468	85149	49554	17994	14924	39650	95294	00556	70481	06905
94559	37559	49678	53119	70312	05682	66986	34099	74474	20740
41615	70360	64114	58660	90850	64618	80620	51790	11436	38072
50273	93113	41794	86861	24781	89683	55411	85667	77535	99892
41396	80504	90670	08289	40902	05069	95083	06783	28102	57816
25807	24260	71529	78920	72682	07385	90726	57166	98884	08583
06170	97965	88302	98041	21443	41808	68984	83620	89747	99882
60808	54444	74412	81105	01176	28838	36421	16489	18059	51061
80940	44893	10408	36222	80562	71944	92638	40333	67054	16067
19516	90120	46759	71643	13177	55292	21036	82808	77501	97427
49386	54480	23604	23554	21785	41101	91178	10174	29420	90438
06312	88940	15995	69321	47458	64809	98189	81851	29651	84215
60942	00307	11897	92674	40405	68032	96717	54244	10701	41393
92329	98932	78284	46347	71209	92061	39448	93136	25722	08564
77936	63574	31384	51924	85561	29671	58137	17820	22751	36518
38101	77756	11657	13897	95889	57067	47648	13885	70669	93406
39641	69457	91339	22502	92613	89719	11947	56203	19324	20504
84054	40455	99396	63680	67667	60631	69181	96845	38525	11600
47468	03577	57649	63266	24700	71594	14004	23153	69249	05747
43321	31370	28977	23896	76479	68562	62342	07589	08899	05985
64281	61826	18555	64937	13173	33365	78851	16499	87064	13075
66847	70495	32350	02985	86716	38746	26313	77463	55387	72681
72461	33230	21529	53424	92581	02262	78438	66276	18396	73538
21032	91050	13058	16218	12470	56500	15292	76139	59526	52113
95362	67011	06651	16136	01016	00857	55018	56374	35824	71708
49712	97380	10404	55452	34030	60726	75211	10271	36633	68424
58275	61764	97586	54716	50259	46345	87195	46092	26787	60939
89514	11788	68224	23417	73959	76145	30342	40277	11049	72049
15472	50669	48139	36732	46874	37088	73465	09819	58869	35220
12120	86124	51247	44302	60883	52109	21437	36786	49226	77837

المصــــدر :

The Rand Corporation, A. Million Random Digits with 100,000 Normal Deviates (Glencoe, III, Free Press, 1955), P. 4, Reprinted with Permission of the Rand Corporation.

جدول (2): المساحات تحت المنحنى الطبيعي

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	0000	0040	0080	0120	0159	0199	0239	0279	0319	0359
0.1	0398	0438	0478	0517	0557	0596	0636	0675	0714	0753
0.2	0793	0832	0871	0910	0948	0987	1026	1064	1103	1141
0.3	1179	1217	1255	1293	1331	1368	1406	1443	1480	1517
0.4	1554	1591	1628	1664	1700	1736	1772	1808	1844	1879
0.5	1915	1950	1985	2019	2054	2088	2123	2157	2190	2224
0.6	2257	2291	2324	2357	2389	2422	2454	2486	2518	2549
0.7	2580	2612	2642	2673	2704	2734	2764	2794	2823	2852
0.8	2881	2910	2939	2967	2995	3023	3051	3078	3106	3133
0.9	3159	3186	3212	3238	3264	3289	3315	3340	3365	3389
1.0	3413	3438	3461	3485	3508	3531	3554	3577	3599	3621
1.1	3643	3665	3686	3718	3729	3749	3770	3790	3810	3830
1.2	3849	3869	3888	3907	3925	3944	3962	3980	3997	4015
1.3	4032	4049	4066	4083	4099	4115	4131	4147	4162	4177
1.4	4192	4207	4222	4236	4251	4265	4279	4292	4306	4319
1.5	4332	4345	4357	4370	4382	4394	4406	4418	4430	4441
1.6	4452	4463	4474	4485	4495	4505	4515	4525	4535	4545
1.7	4554	4564	4573	4582	4591	4599	4608	4616	4625	4633
1.8	4641	4649	4656	4664	4671	4678	4686	4693	4699	4706
1.9	4713	4719	4726	4832	4738	4744	4750	4758	4762	4767
2.0	4773	4778	4783	4788	4793	4798	4803	4808	4812	4817
2.1	4821	4826	4830	4834	4838	4842	4846	4850	4854	4857
2.2	4861	4865	4868	4871	4875	4878	4881	4884	4887	4890
2.3	4893	4896	4898	4901	4904	4906	4909	4911	4913	4916
2.4	4918	4920	4922	4925	4927	4929	4931	4932	4934	4936
2.5	4938	4940	4941	4943	4945	4946	4948	4949	4951	4952
2.6	4953	4955	4956	4957	4959	4960	4961	4962	4963	4964
2.7	4965	4966	4967	4968	4969	4970	4971	4972	4973	4974
2.8	4974	4975	4976	4977	4977	4978	4879	4980	4980	4981
2.9	4981	4982	4983	4984	4984	4984	4985	4985	4986	4986

x/σ	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
3.0	4986.5	4987	4987	4988	4988	4988	4989	4989	4989	4990
3.1	4990.3	4991	4991	4991	4992	4992	4992	4992	4993	4993
3.2	4993.129									
3.3	4995.166									
3.4	4996.631									
3.5	4997.674									
3.6	4998.409									
3.7	4998.922									
3.8	4999.277									
3.9	4999.519									
4.0	4999.683									
4.5	4999.966									
5.0	4999.997133									

إيضـــاح :

أجزاء الكسور المستخدمة في العدد الكلي (10,000) تحت المنحنى الطبيعي الاحتمالي، تطابق في قيمتهـا، مسـافـات الخط الـرئيسي الـذي يصل مابين المتوسط الحسابي والنقاط المتتابعة للأجزاء المستقطعة من المتوسط الحسابي، وهذه المسافات مقاسة بوحدات الانحراف المعياري (σ) ويقرأ الجـدول كالآتي: بين المتـوسط الحســابي للإحداث الرأسي (y) ، وأي إحداثي موجود على مسافة منه، ولتكن 8 σ (مثال = .8 = $\frac{x}{\sigma}$)، مُضَمَن في نسبة الـ 28,81% من المساحة الكلية .

Harold O. Rugg, Statistical methods applied to Education (New York: Houghton Mifflin 1917), p. 389-90, Copyright 1917. by Houghton Mifflin company, Used with permission.

جدول (3): توزيع t

df	Level of significance for one-tailed test					
	.10	.05	.025	.01	.005	.0005
	Level of significance for two-tailed test					
	.20	.10	.05	.02	.01	.001
1	3.078	6.314	12.706	31.821	63.657	636.619
2	1.886	2.920	4.303	6.965	9.925	31.598
3	1.638	2.353	3.182	4.541	5.841	12.941
4	1.533	2.132	2.776	3.747	4.604	8.610
5	1.476	2.015	2.571	3.365	4.032	6.859
6	1.440	1.943	2.447	3.143	3.707	5.959
7	1.415	1.895	2.365	2.998	3.499	5.405
8	1.397	1.860	2.306	2.896	3.355	5.041
9	1.383	1.833	2.262	2.821	3.250	4.781
10	1.372	1.812	2.228	2.764	3.169	4.587
11	1.363	1.796	2.201	2.718	3.106	4.437
12	1.356	1.782	2.179	2.681	3.055	4.318
13	1.350	1.771	2.160	2.650	3.012	4.221
14	1.345	1.761	2.145	2.624	2.977	4.140
15	1.341	1.753	2.131	2.602	2.947	4.073
16	1.337	1.746	2.120	2.583	2.921	4.015
17	1.333	1.740	2.110	2.567	2.898	3.965
18	1.330	1.734	2.101	2.552	2.878	3.922
19	1.328	1.729	2.093	2.539	2.861	3.883
20	1.325	1.725	2.086	2.528	2.845	3.850
21	1.323	1.721	2.080	2.518	2.831	3.819
22	1.321	1.717	2.074	2.508	2.819	3.792
23	1.319	1.714	2.069	2.500	2.807	3.767
24	1.318	1.711	2.064	2.492	2.797	3.745
25	1.316	1.708	2.060	2.485	2.787	3.725
26	1.315	1.706	2.056	2.479	2.779	3.707
27	1.314	1.703	2.052	2.473	2.771	3.690
28	1.313	1.701	2.048	2.467	2.763	3.674
29	1.311	1.699	2.045	2.462	2.756	3.659
30	1.310	1.697	2.042	2.457	2.750	3.646
40	1.303	1.684	2.021	2.423	2.704	3.551
60	1.296	1.671	2.000	2.390	2.660	3.460
120	1.289	1.658	1.980	2.358	2.617	3.373
∞	1.282	1.645	1.960	2.326	2.576	3.291

المصــدر :

Table III of Fisher and Yates, Statistical tables .. for Biological, Agricultural, and Medical Research, Published by Longman. Group Ltd, London (Previosly Published by Oliver and Boyed Edinbargh), Used with Permission of authors and publishers.

جدول (4): توزيع مربع كاي (X^2 : كا ²)

df	.99	.98	.95	.90	.80	.70	.50	.30	.20	.10	.05	.02	.01	.001
1	$.0^3157$	$.0^3628$	.00393	.0158	.0642	.148	.455	1.074	1.642	2.706	3.841	5.412	6.635	10.827
2	.0201	.0404	.103	.211	.446	.713	1.386	2.408	3.219	4.605	5.991	7.824	9.210	13.815
3	.115	.185	.352	.584	1.005	1.424	2.366	3.665	4.642	6.251	7.815	9.837	11.341	16.268
4	.297	.429	.711	1.064	1.649	2.195	3.357	4.878	5.989	7.779	9.488	11.668	13.277	18.465
5	.554	.752	1.145	1.610	2.343	3.000	4.351	6.064	7.289	9.236	11.070	13.388	15.086	20.517
6	.872	1.134	1.635	2.204	3.070	3.828	5.348	7.231	8.558	10.645	12.592	15.033	16.812	22.457
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337	28.172	30.675	34.382	37.652	41.566	44.314	52.620
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	12.879	14.125	16.151	18.114	20.703	22.719	26.336	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336	32.461	35.139	39.087	42.557	46.693	49.588	58.302
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336	33.530	36.250	40.256	43.773	47.962	50.892	59.703

إيضـــاح :

للقيم الأكبر لدرجة الحرية (df) ، تستخدم المعادلة التالية :

١ - $\sqrt{2x^2} - \sqrt{2df}$، كانحراف طبيعي مع وحدة التغاير، مع الأخذ في الاعتبار، أن قيمة مربع كاي المحتملة تطابق قيمة طرفية واحدة من المنحنى الطبيعي.

Table IV of Fisher and Yates, Statistical tables .. for Biological, Agricultural, and Medical Research, Published by Longman. Group Ltd, London (Previously Published by Oliver and Boyed Edinburgh), Used with Permission of authors and publishers.

جدول (5): قيم معامل الارتباط، للمستويات المختلفة للدلالة الإحصائية

df	*P* = .10	.05	.02	.01
1	.988	.997	.9995	.9999
2	.900	.950	.980	.990
3	.805	.878	.934	.959
4	.729	.811	.882	.917
5	.669	.754	.833	.874
6	.622	.707	.789	.834
7	.582	.666	.750	.798
8	.549	.632	.716	.765
9	.521	.602	.685	.735
10	.497	.576	.658	.708
11	.476	.553	.634	.684
12	.458	.532	.612	.661
13	.441	.514	.592	.641
14	.426	.497	.574	.623
15	.412	.482	.558	.606
16	.400	.468	.542	.590
17	.389	.456	.528	.575
18	.378	.444	.516	.561
19	.369	.433	.503	.549
20	.360	.423	.492	.537
21	.352	.413	.482	.526
22	.344	.404	.472	.515
23	.337	.396	.462	.505
24	.330	.388	.453	.496
25	.323	.381	.445	.487
26	.317	.374	.437	.479
27	.311	.367	.430	.471
28	.306	.361	.423	.463
29	.301	.355	.416	.456
30	.296	.349	.409	.449
35	.275	.325	.381	.418
40	.257	.304	.358	.393
45	.243	.288	.338	.372
50	.231	.273	.322	.354
60	.211	.250	.295	.325
70	.195	.232	.274	.302
80	.183	.217	.256	.283
90	.173	.205	.242	.267
100	.164	.195	.230	.254

إيضـــاح :

الاحتمالات المُعطاة، تخص اختبار طرفين من الدالة الإحصائية، وذلك عندما تُهمل إشارة r ، وفي حالة إجراء اختبار طرفي واحد للدلالة الإحصائية، لتقييم الاحتمالات المجدولة يجب أن تقسم على 2 (إلى النصف).

Table VII of Fisher and Yates, Statistical Tables .. for Biological, Agricultural, and Medical Research, Published by Longman. Group Ltd, London (Previosly Published by Oliver and Boyed Edinbargh).

جدول (6): القيم الحرجة لـ R ، في اختيار الـ Runs (حيث P = ,05)

N_1 \ N_2	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
4			2																
5		2	2	3															
6		2	3	3	3														
7		2	3	3	4	4													
8	2	2	3	3	4	4	5												
9	2	2	3	4	4	5	5	6											
10	2	3	3	4	5	5	6	6	6										
11	2	3	3	4	5	5	6	6	7	7									
12	2	3	4	4	5	6	6	7	7	8	8								
13	2	3	4	4	5	6	6	7	8	8	9	9							
14	2	3	4	5	5	6	7	7	8	8	9	9	10						
15	2	3	4	5	6	6	7	8	8	9	9	10	10	11					
16	2	3	4	5	6	6	7	8	8	9	10	10	11	11	11				
17	2	3	4	5	6	7	7	8	9	9	10	10	11	11	12	12			
18	2	3	4	5	6	7	8	8	9	10	10	11	11	12	12	13	13		
19	2	3	4	5	6	7	8	8	9	10	10	11	12	12	13	13	14	14	..
20	2	3	4	5	6	7	8	9	9	10	11	11	12	12	13	13	14	14	15

إيضـــاح :

لأي اختبار لـ Runs لعينتين، أي قيمة لـ R تكون مساوية أو أقل من القيمة الموضحة في الجدول، تُعد ذات دلالة إحصائية، في المستوى الدلالي 05, إذا كانت قيمة R غير متنبأ بها. أي قيمة لـ R مساوية أو أقل من القيمة الموضحة في الجدول، تُعد ذات دلالة إحصائية في المستوى الدلالي 025, إذا ما كانت قيمة R غير متنبأ بها.

المصــدر :

F.S. Swed and C. Eisenhart, "Tables for Testing Randomeness of Grouping in Sequence of Alternatives", Annals of Mathematical Statistics, 14:83-86 (1943). Reprinted with permission.

GLOSSARY قائمة مصطلحات

القيمة المطلقة **Absolute Value**

القيمة المطلقة لـ X ، هي المسافة بين X و O ، أو القيمة الإيجابية لـ X. والقيمة المطلقة لـ (23-) = 23, والقيمة المطلقة لـ (5) = 5 ويشار إلى القيمة المطلقة بالصيغة التالية:

$$13 = |13 -|$$

عينة الصدفة **Accidental Sample**

من أكثر العينات سهولة ويسراً في الحصول عليها وإيجادها.

التوزيع ثنائي النموذج **Bimodal Distribution**

توزيع حالات مُقاس على ضوء متغير واحد، كما يحدث عندما تظهر قيمتان لمتغير واحد، في المجموعة المقاسة، بنفس القدر من التكرار، وذلك عندما يكون هذا الظهور المتكرر هو الأكبر حجماً للمجموعة المقاسة، ويُعد هذا بمثابة منوالين أو قيمالظهور الثنائية الأكثر تكراراً.

حالة بحثية (دراسية) **Case**

هي الوحدة التي تقاس بها المتغيرات. أمثلة: إذا قمنا بقياس عدد المجلدات في مجموعة مكتبات، فالمكتبات ـ هنا ـ تمثل الحالات البحثية. وإذا ما قمنا بقياس المرتبات لمجموعة من المكتبيين، فالمكتبيون هم الذين يمثلون الحالات البحثية.

معامل الانحـــراف **Coefficient of Variation**

هو نسبة الانحراف المعياري، عن المتوسط الحسابي للمجموعات البحثية، المقاسة على ضوء متغير واحد في المجموعة. ويفيد ـ معامل الانحراف ـ عندما نرغب في إجراء مقارنة مابين مجموعتين بحثيتين أو أكثر، مقاسة على ضوء متغير واحد (أي نفس المتغير).

معامل الثقــة Confidence Coefficient

يرتبط مستوى الاحتمال ـ بوجه عام ـ بالحد. وتشير القيمة 99 (95, أو 90,) إلى أننا لو كررنا إجراءات المعاينة الإحصائية، لعدد غير محدود من المرات، فلنا أن نتوقع أن قيمة المتـوسط الحسـابي لمجتمع الدراسة سوف يقع في كل مرة، داخل حد الـ 99% التي حددناها. وهذا يعني أن نتوقع في 99% من المرات، أن يقع المتوسط الحسابي للعينة، داخل الحد (حد الثقة) الذي قمنا بتحديده.

حــد الثقــة Confidence Interival

وهــو الخط المستقيم، أو الحــد الفـاصـل، الـذي تُسجل عليه معلمات المجتمع الدراسي، والإحصاءات الأخرى للعينة.

حــدود الثقـة Confidence Limits

القيم الخارجية لحد الثقة لو كان حد الثقة يعني بالمتوسطات الحسابية (Confidence Coefficient) فالقيم الخارجية تسـاوى قيم، ($\bar{X}$) (المتوسط الحسابي).

جدول الجدولة ـ المتقاطعة Cross-Tabulation Table

أبسط أشكاله، جدول (2x2)، حيث تشير فيه، قيم الأعمدة إلى القيم المحتملة لمتغير واحد، وتشير فيه الصفوف إلى القيم المحتملة للمتغير الآخر. وتصنف فيه مجموعة واحدة من الحالات البحثية للمتغيرين، وكل حالة توضع داخل خلية واحدة، تمثل نقطة تقاطع واحد من الصفوف مع واحد من الأعمدة، ويسجل العدد الكلي للحالات للخلية الواحدة، ويجدول داخل كل خلية (أي تسجل قيمته داخل كل خلية).

درجـات الحرية Degrees of Freedom

تُحسب درجة الحرية، في اختبار t، عن طريق حساب (جمع) الحجم الكلي للعينة، ونطرح منه حجم الحالات الدراسية معروفة القيمة في العينة، والعدد الناتج عن عملية الـطرح هو درجات الحرية (أو حجم الحالات الدراسية مجهولة الحجم) ويمثل عدد القيم التي تتميز بحرية التغاير في قيمتها. مثال (أنظر: «اختبار الفروض حول (μ) عنـدمـا تكون (n) صغيرة الحجم». الفصل الثالث: الإحصاء الاستنباطي، الفقرة الثالثة). أما في اختبار مربع كاي، فتكون قيمة درجة الحرية مساوية لعدد الصفوف ناقصاً واحد، مضروباً في عدد الأعمدة ناقصاً واحد: (1-r)(1-c).

المتغير التابع **Dependent Variable**

المتغير المطلوب اختباره وقياسه.

الإحصاء الوصفي **Descriptive Statistics**

هي الإحصاءات المستخدمة لتلخيص معلومات، حول مجموعة حالات بحثية، مُقاسة على ضوء متغير أو أكثر، ولا ينتج عن بياناتها أي نوع من أنواع الإستنتاج أو الإستنباط، ونستخدمها في وصف النزعة المركزية، ودرجة تغاير كل متغير لمجموعة حالات البحث، وتستخدم ـ أيضاً ـ في تحديد العلاقات الكائنة بين متغيرات عينة البحث.

الخطأ المعياري التقديري **Estimated Standard Error**

هو الانحـراف المعياري لتوزيع المعاينة الإحصائية، يُحسب عن طريق استخدام الانحراف المعياري للعينة، بدلاً عن الانحراف المعياري لمجتمع الدراسة، وبما أننا نعتمد في تقديرنا على كلية مجتمع الدراسة لتحديد الانحراف المعياري، فقـيم النتائج تُعد خطـأ معياري متوقع (أي خطا معياري تقديري).

التكرار المتوقع **Expected Frequency**

عدد المفردات المتوقعة في الخلية في جدول الجدولة المتقاطعة، عندما يكون المتغيران موضع الدراسة مستقلين تماماً عن بعضهما.

توزيع التكرار **Frequency Dstribution**

جدول يحتـوي على عمودين: قيم المتغير، والتكرار. وقيم المتغير، تعني، القيمة المحتملة التي يأخذها المتغير في عينة البحث، ويشير التكرار إلى عدد الحالات البحثية التي ظهر مقترنة بقيمة المتغير للمرة الأولى، زائداً عدد الحالات التي تظهر بقيمة المتغير للمرة الثانية... وهكذا.

المضلع التكراي **Frequency Polygon**

رسم بياني ذو بعدين للتوزيع التكراري، حيث يشير فيه الخط البياني الأفقي إلى قيم المتغـير، ويشير فيه الخط البياني الرأسي إلى التكرار أو عدد الحالات البحثية في العينة، ويستخدم عادة لقياس البيانات ـ على الأقل ـ في المستوى الترتيبي للقياس.

المدرج التكراري **Histogram**

رسم بياني عمودي، يستخدم لتوضيح التوزيع التكراري، حيث يشير فيه الخط

البياني الأفقي إلى قيم المتغيرات، والخط البياني الرأسي إلى تكرار ظهور القيم في العينة، ويمكن استخدام المدرج التكراري، لقياس متغيرات في المستوى الإسمي للقياس.

اختبار الفروض **Hypothesis Test**

عبارة عن اختبار لفروض إحصائية محددة، تختص بمقاييس مجتمع الدراسة، تتم عن طريق معايير توزيع المعاينة الإحصائية الموضوعة على ضوء الفروض الإحصائية، وتحديد حجم العينة، ثم اختيار عينة عشوائية من مجتمع البحث موضع الدراسة. ويُجري بعد ذلك مقارنة الإحصاءات المعنية من المعاينة الإحصائية (الموضوعة على ضوء الفروض الإحصائية) وذلك بغرض تحديد ملائمتها أو عدم ملاءمتها للبحث وفروضه، وبالتالي نأخذ قرار (رفض) أو قبول الفرض الإحصائي (الفرض الصفري).

المتغير المستقل **Independant Variable**

المتغير المُعالج بغرض اختبار تأثيره على المتغير التابع، وغالباً ما يسبق المتغير التابع في ظهوره.

الإحصاء الاستنباطي (الاستدلالي) **Inductive Statistics**

أنظر الإحصاء الاستنتاجي (Inferential Statistics)

الإحصاء الاستنتاجي **Inferential Statistics**

هو ذلك النوع من الإحصاءات التي يطبق فيها منطق الاستنباط، بغرض تقييم النتـائج عن المجتمع الدراسي ككل، اعتماداً على معلومات جُمعت عن طريق العينة العشوائية للحالات الدراسية (محاولة تطبيق نتائج الجزء على الكل).

التقاطع (قيمة نقاط الانحدار) **Intercept**

يمثل قيمة (a) «درجة الانحدار»، في معادلة الانحدار التي تتعلق بمتغير واحد (X)، حيث تتقاطع الخطوط مع الخط الرأسي (Y) في الرسم البياني.

الحـــــد **Interval**

مستوى قياس، حيث تُعطي للحالات البحثية أعداد، بناء على قواعد مستويات القياس الترتيبي والإسمي، بالاضافة إلى وجود حدة قياس ثابتة، بمعنى أن المسافة بين 5 و 6 تأخذ نفس قيمة المسافة بين 9,8 وبناء على قيمة المتغير موضع القياس مع افتراض وجود صفر عشوائي. مثال: يمكن أن يقاس ذكاء الأفراد على ضوء المتغير IQ أحد أنواع اختبارات الذكاء).

المستوى الدلالي الإحصائي **Level of Significance**

تمثله نسبة المساحة الواقعة تحت توزيع المعاينة الإحصائية، والذي يمثل ما يمكن اعتباره منطقة الرفض، أو القيم غير المحتمل ظهورها بناء على افتراضنا الصفري. ويُعد المستوى الدلالي، هو الاحتمال بأن نحصل على قيمة غير محتملة الظهور، أو أن احتمال ظهورها أقل من احتمال ظهور القيم الواقعة في منطقة الرفض.

مستوى القياس **Level of Measurement**

عادة يوجد أربع مستويات للقياس: القياس الإسمي، القياس الترتيبي، القياس الحدي، والقياس النسبي. وفيها تتطابق قيم مجموعة الحالات البحثية. المُقاسة على ضوء متغير واحد مع بعضها البعض، بناء على خواص رياضية محددة.

المتوسط الحسابي (الوسط الحسابي) **Mean**

أحد أنواع مقاييس النزعة المركزية، يُعطي عدداً واحداً لمجموعة مُقاسة على ضوء متغير واحد. ويشير المتوسط الحسابي إلى متوسط قيمة المتغير لمجموعة حالات بحثية، ويجب أن يتم حسابه على بيانات مقياس مستوى حدى ـ على الأقل ـ، ويجري حسابه بالقيم الطرفية (الأعلى والأدنى)، وبالارتباط مع الانحراف المعياري، وباستخدام هذه المعادلة:

$$\bar{X} = \sum_{i=1}^{n} \frac{X_i f_i}{n}$$

الانحراف المطلق للمتوسط الحسابي **Mean Absolute Deviation**

يعني، قياس التغـــاير للوسط الحســابي (أي انحـرافـه عن المعـدل الـطبيعي)، باستخدام متغير واحد، بغرض بيان تجانس أو عدم تجانس مجموعة الحالات البحثية (العينـة) الخـاصـة بالبحث. ولكنـه ليس في مستـوى فائـدة الانحـراف المعياري، كمقياس للتغاير، حيث أنه لا يشتمل على نفس العمليات الحسابية المتعلقة بالتوزيع الطبيعي.

القياس **Measurement**

إجراء إحصائي / رياضي، يتم عن طريقة تحديد أعداد لتمثل المستويـات المختلفة التي يمكن للمتغير أن يحـوز عليها. مثـال: المتغير «الجنس»، يمكن أن يُعطـي القيمة = 5 للرجال والقيمة = 1 للنساء.

المتوسط العددي (الوسط العددي) Median

أحد أنواع مقاييس النزعة المركزية. التي لا تحسب عن طريق القيم الطرفية، وهو قيمة المتغير بالارتباط مع الحالات البحثية، حيث يسبقه 50% ويعقبه 50% من مفردات العينة (أو مفردات مجتمع الدراسة)، ويتم حساب المتوسط العددي للمتغيرات المقاسة، على بيانات مقياس مستوى ترتيبي، على الأقل.

المنـوال Mode

أحد أنواع مقاييس النزعة المركزية، يشير إلى القيمة الأكثر تكراراً في الظهور، لمتغير مقاس في مجموعة واحدة من الحالات البحثية.

الرمـز n

يستخدم لبلاشارة إلى عدد المشاهدات، أو عدد الحالات البحثية، في عينة بيانات. مثال: لو قمنا بقياس 45 مكتبة، على ضوء 30 متغير خاضع للدراسة، فحجم العينة أو n = 45 (يشار إليه في الترجمات بـ ن).

الرمز N

يشير إلى العدد الكلي للمشاهدات لعينات مجتمع دراسي: مثال:

إذا ما قمنا بقياس عينة تتشكل من 45 فرداً على ضوء 30 متغيراً، فحجم العينة n = 45، ولكن حجم N، يساوي العدد الكلي للمكتبيين المراد تعميم النتائج عليهم، مثال لذلك عدد المكتبيين العاملين في الولايات المتحدة الأمريكية [وهذا يعني: إذا كان n = حجم العينة (مفردات العينة)، N = حجم مجتمع البحث المسحوبة منه هذه العينة (عدد مفرادات مجتمع الدراسة بالكامل)].

أسمى (مستوى القياس الإسمي) Nominal

مستوى قياس، وفيه يلحق مفرد أو حالة بحث إلى نوع معين من المتغيرات، ولا يتطلب هذا القياس، وجود أي ترتيب في عرض البيانات. مثال: نظام ديوي العشري، الجنس، الانتماء إلى الأحزاب السياسية... الخ. يُعد العدد الذي يلحق بالمتغير «كحافظ للترتيب المكاني فقط»، ولا يتطلب إجراء أي عمليات حسابية ذات معنى. مثال: الجنس: 1 = ذكر، 2 = أنثى. فلو حصل الشخص على رقم (1) في متغير الجنس، فهذا يعني أنه ذكر، بمعنى أن هذا الرقم لا يشير إلى أي ترتيب نظامي فيما يتعلق بالجنس [وهذا يعني، أن القياس الإسمي، لا يتضمن أي

دلالة إحصائية/ حسابية مطلقاً، فهو مجرد تخصيص أرقام المفردات، لكي نتمكن من وضعهـا على قوائم، ولا تمثل هذه الأرقام أي دلالة أو قيمة أو مغزى إحصائي، أو حسابي أو مفهومي أو تصنيفي، فهو بالكاد مؤشر لمكان وجود المفـردات على القائمة. (المترجمان)].

المعاينة الإحصائية غير الإحتمالية **Non-Probability Sampling**

نوع من المعـاينة الإحصائية، حيث كل مفرد من مفردات مجتمع الدراسة، لديه فرصة احتمال معروفة (احتمال مجهول)، لِيُضمن في العينة المسحوبة من هذا المجتمع.

التوزيع الطبيعي **Normal Distribution**

عائلة (مجموعة) من المنحنيات البيانية أو التوزيعات الرياضية، التي تُعد في غاية الأهمية في مجال البحوث، طالما أن العديد من المتغيرات تتبع التوزيع الطبيعي في مجتمع الدراسة. يتبع المتغير التوزيع الطبيعي، لوكان توزيع التكرار لهذا المتغير يأخذ شكل التوزيع الطبيعي. ويشار ـ أحياناً ـ إلى التوزيع الطبيعي بـ «منحنيات شكل الجرس»، وقـد تم شرح معـطيات هذه المنحنيات في فصل (1) في حالة اتخاذ المتغير أو المعاينة الإحصـائية للمتـوسط الحسابي شكل التوزيع الطبيعي، يمكننا إيجاد النسب المئوية للحالات الدراسية التي تقع بين أي متغيرين موضع اهتمامنا، وذلك باستخدام جدول رقم (2) في الملحق.

الفرض الصفري **Null Hypothesis**

أنظر الفرض الإحصائي Statistical Hypothesis

ترتيبي (مستوى القياس الترتيبي) **Ordinal**

مستوى قياس، وفيه يلحق مفرداً أو حالة بحث إلى نوع معين من المتغيرات، وترتب فيه، بنظام معين، مثال: قد يقوم مستفيد بترتيب 10 كتب، حسب الأفضلية التي يراها مناسبة من وجهة نظره، في هذا المثال، تمثل الـ 10 كتب، 10 حالات دراسية، والمتغير هو نظام الترتيب بالأفضليات الذي اتبعه المستفيد.

معلمـــة **Parameter**

الخصائص الكمية لمجتمع الدراسة، وهي قيمة المتغير التي يمكن الحصول عليها، إذا كانت كل مفردات البحث مقاسة على ضوء المتغير موضع الدراسة، وهي الرقم الثابت الذي ـ عادة ـ يكون مجهولا (الثابت الذي يغير قيمته حسب المجال الذي ينتمي إليه).

معامل الارتباط الخطي لبيرسون (معامل ارتباط بيرسون)

Pearson Product – Moment Correlation Coefficient

مقياس لقوة واتجاه العلاقة الخطية بين متغيرين في عينة حالات بحثية [لقياس شدة الارتباط/ واتجاهه (سالب/ موجب)].

النسبة المئويـة **Perecentage**

مقدار (قيمة) مضروب في 100 ـ مثال: 77% تساوي 77x100, أو 77%. ويشير إلى 77% من العدد الكلي للحالات البحثية، أو إلى 77% من القيمة الكلية.

مجتمع الدراسـة **Population**

العدد الكلي للحالات ا لدراسية موضع البحث، أو العدد الكلي للحالات الدراسية التي يمكن تعميم نتائجها ومن الممكن أن يكون المجتمع الدراسي قابل للعد. أو من الكبر بمكان بحيث تكون عدد مفرداته لا نهائية، ولذلك، يمكن أن نُعده غير قابل للعد.

توزيع مجتمع الدراسة **Population Distribution**

يعني توزيع التكرار لمجتمع حالات دراسية، على ضوء متغير محدد موضع دراسة. وبما أننا ـ نادراً ـ ما يكون لدينا القدرةعلى دراسة جميع مفردات مجتمع الدراسة، لذا يُعد توزيع مجتمع الدراسة ـ بوجه عام ـ توزيعاً نظرياً.

المعاينة الإحصائية الاحتمالية **Probability Sampling**

نوع من العينة، تتميز بأن كل مفرد من مفردات المجتمع الدراسي الذي سُحبت منه، يحتمل اختياره في العينة المسحوبة، وهو احتمال معلوم مسبقاً.

التناســب **Proportion**

كسر، يعبر عنه عادة في صورة عشرية، وفيه يتم المقارنة ما بين عدد صغير من الحالات البحثية والعدد الكلي للحالات. مثال: عدد العاملين في مكتبة ما = 30، منهم 20 إمرأة و 10 رجال، إذن، نسبة النساء إلى المجموع الكلي = $\frac{20}{30}$ أو 667,.

العينة القصدية (الغرضية) **Purposive Sample**

عينـة غير احتمالية، وفيها يقوم الباحث باختيار الحالات البحثية، اعتماداً على حكمه

ورأيه الشخصي. مثال: اختيار الباحث لخبراء في موضوع بحثه، وتشكيل عينة منهم.

الحصة Quata

عينة غير احتمالية [يطلق عليها Quata Sample العينة الحصصية، وفيها يقوم الباحث باختيار الحالات البحثية بناء على أعداد مسبق لعدد مفردات العينة (حجم العينة)، وتحديد الخصائص المميزة لكل مفرد من المفردات، بحيث تتضمن خصائص المجتمع الدراسي في العينة المختارة. وبذلك تكون العينة ممثلة المجتمع الدراسي المسحوبة منه.

المدى Range

أقل أنواع قياسات التغاير فائدة، يحسب عن طريق طرح أصغر قيمة تظهر للمتغير من أكبر قيمة له. مثال: لنفرض في عينة ما، كانت القيمة الصغرى للمتغير A = 20 وأكبر قيمة ظهرت للمتغير = 50 ، يكون المدى = 50 − 20 = 30.

النسبة (مستوى القياس النسبي) Ratio

مستوى قياس، يتضمن كل قيم مستويات القياس الإسمي، والترتيبي، والحدي، بالاضافة إلى وجود الصفر المطلق. الذي يشير إلى الغياب الكلي لقيمة المتغير في العينة المقاسة. مثال: عدد الكتب في مكتبة ما، مرتبات العاملين في مكتبة ما... الخ.

النسبة Ratio

نسبة X إلى y تساوي عدد الحالات البحثية لـ X مقارنة مع عدد الحالات البحثية لـ y. مثال: لو لدينا مجموعة بحثية مكونة من 10 ذكور، 15 نساء، فإن نسبة الذكور إلى النساء تساوي 10 إلى 15 ويعبر عنها بالشكل التالي: 15:10 أو 3:2.

تحليل الانحدار Regression Analysis

تقنية إحصائية تتضمن واحداً أو إثنين من المتغيرات المستقلة (X_1X_2)، ومتغير واحد غير مستقل (تابع) (y)، وفيها يمكن تحديد شكل العلاقة بين المتغيرات، والأهمية النسبية لكل من المغيرات X_i (المستقلة)، للمساعدة في التنبؤ بقيمة (y) (المتغير التابع).

منطقة الرفض Rejection Region

في توزيع المعاينة الإحصائية، تتضمن هذه المنطقة العينات الإحصائية غير محتملة الظهور، بناء على الفرض الصفري، وعموماً تعني المناطق الطرفية، أو قيم العينة التي تظهر نادراً، لو كان الفرض الإحصائي (الصفري) صحيحاً.

عينـة **Sample**

مجموعة فرعية من العدد الكلي للحالات البحثية، أو مجموعة فرعية من حالات مجتمع الدراسة.

توزيع العينـة **Sample Distribution**

توزيع التكرار لعينة الحالات الدراسية، على ضوء متغير محدد.

حجم العينـة **Sample Size**

عدد الحالات الدراسية أو المشاهدات في العينة، وتعني عدد مفردات البحث المطلوب قياس متغيراتها.

توزيع المعاينة الإحصائية **Sampling Distribution**

توزيع نظري لكل القيم المحتملة الظهور للعينة الإحصائية، بشرط صحة الفرض الصفري، حيث تستند الإحصاءات على عينة عشوائية مستقلة بحجم محدد (n)، مسحوبة جميع مفرداتها من نفس المجتمع الدراسي.

شكل الانتشار **Scattergram**

رسم بياني يوضح كمية «الارتباط» بين متغيرين، حيث يمثل الخط البياني الأفقي قيم متغير واحد (المتغير الأول)، والخط البياني الرأسي يمثل قيم المتغير الثاني. وتشير النقاط على الرسم البياني (الخط البياني) إلى قيمتي الحالة الواحدة، مقاسة على ضوء المتغيرين.

المعاينـة العشوائية البسيطة **Simple Random Sampling**

العينة العشوائية البسيطة، نوع من العينات الاحتمالية، وفيها يكون كل مفرد من مفردات مجتمع الدراسة، لدية فرص متكافئة ليُضمن في العينة المختار عشوائياً.

الانحدار (قيمة الانحدار) **Slope: (b)**

في معادلة انحدار لمتغيرين، فإن قيمة (b) تساوي انحدار أو هبوط خط الانحدار عندما يُرسم بيانياً.

الانحراف المعياري **Standard Deviation**

مقياس تغير لمتغير واحد. يشير إلى تجانس مجموعة الحالات الدراسية، على ضوء المتوسط الحسابي للمجموعة. ويشير الانحراف المعياري الكبير القيمة، إلى أن

الفروقات بين مفردات البحث كبيرة. ويشير الانحراف المعياري الصغير القيمة، إلى أن المجموعة متجانسة أو متشابهة ومتعنقدة بشدة حول المتوسط الحسابي للمجموعة. ويعد مقياساً مفيداً للغاية، عندما يتبع المتغير توزيعاً طبيعياً.

الخطأ المعياري **Standard Error**

يعني الانحراف المعياري لتوزيع عينة إحصائية، مفحوصة لأغراض استنباطية. أمثلة: الخطأ المعياري للمتوسط الحسابي، الخطأ المعياري للتباين (المتغاير)... الخ.

إحصاء / إحصاءات **Statistics**

الخواص الكمية للعينة. أمثلة المتوسط الحسابي للعينة، rs للعينة، الانحراف المعياري للعينة، المنوال، .... الخ. والإحصاءات تختلف قيمتها من عينة إلى أخرى.

الفرض الإحصائي **Statistical Hypothesis**

يطلق عليه أيضاً «الفرض الصفري»، وهو بيان رياضي حول قيم معلمه (أو معلمات) لمجتمع الدراسة المحدد، موضع اهتمامنا. أمثلة: $\mu = 30$، $P = 70$، $t = 150$، ويفترض صحة الفرض الإحصائي (الصفري) حتى يثبت العكس من خلال تحليل البيانات.

الدلالة الإحصائية (دالة إحصائية) **Statistical Significant**

تُعد العينة الإحصائية، ذات دلالة إحصائية، وذلك عندما تقع داخل منطقة الرفض في حالة اختبار الفرض الصفري. ويرفض الفرض الصفري، طالما أن العينة الإحصائية تقع بعيداً عن معلمتنا المفترضة.

المعاينة الإحصائية الطبقية **Stratified Sampling**

ويطلق عليها «العينة الطبقية»، وهو نوع من العينات الإحصائية الاحتمالية، وفيها يقسم مجتمع الدراسة إلى عدد من الفئات المختلفة، تُختار فيها الفئات أو الحالات الدراسية عشوئياً. (سواء بصورة متناسبة أو غير متناسبة)، من داخل الفئات التي تمثل مجتمع الدراسة. والطبقة (الفئة) المختارة يجب أن تحتوي على المتغير المناسب والمطلوب للدراسة، ولهذا نقوم بتضمين الحالات الدراسية من كل مستوى من مستويات المتغير.

المعاينة الإحصائية المنتظمة **Systematic Sampling**

ويطلق عليها «العينة المنتظمة»، وهو نوع من العينات الاحتمالية، وفيها يتم اختيار

المفرد الأول (الحالة الأولى) من مجتمع الدراسة عشوائياً، وتختار المفردات أو الحالات البحثية التالية عن طريق مسافات أو فواصل ثابتة، يرمز إليها بالرمز K.

إختبــار t t Test

ويعرف أيضا بتوزيع ف، وهو اختبار إحصائي يتضمن فرض حول المتوسط الحسابي لمتغير مجتمع الدراسة، أو فرض حول الفروقات بين المتوسطان الحسابيان لمجتمع الدراسة، وعموماً، يكون حجم العينة (n) لاختبار t صغيراً، وفي حالة أن يكون حجم العينة صغيراً جداً، (حيث n = أقل من 30)، يجب افتراض أن المتغير في مجتمع الدراسة موزع توزيعاً طبيعياً.

عدم التحيز الإحصائي **Unbiased Distribution**

تُعد الإحصاءات غير متحيزة، عندما يكون المتوسط الحسابي لتوزيع المعاينة الإحصائية مساوياً للمعلمة التي تم تقديرها.

التوزيع أحادي النموذج **Unimodal Distribution**

عبارة عن، توزيع الحالات الدراسية، على ضوء متغير واحد، مثلا، في حالة أن مجموعة الحالات الدراسية تحتوي على منوال واحد أو قيمة واحدة للمتغير، تظهر بشكل مستمر.

المتغـــير **Variable**

أي بند، خاصية، أو ظاهرة تخضع للتحليل ويقاس المتغير للمفرد من مفردات البحث أو أي حالة بحثية، بناء على قواعد نقوم بتحديدها. ويشير مصطلح «متغير» إلى أن القيم تختلف من حالة بحثية إلى أخرى. مثال الجنس، اختبار الذكاء IQ، العمر، عدد الكتب، ترتيب الافضليات . . . الخ . .

التغـاير / التبايـن **Variance**

مربع قيمة الانحراف المعياري.

الدرجة المعيارية **Z Score**

وحدة لقياس الانحراف المعياري، وتعني التحول الذي يطرأ على المتغير، بمعدل حالة بحثية واحدة في كل مرة، وذلك فيما إذا كانت قيمة المتوسط الحسابي للحالات تساوي صفر، والانحراف المعياري يساوي واحد صحيح. ولو كان توزيع المتغيرات طبيعياً، فبإمكاننا استخدام وحدة التوزيع الطبيعي لإيجاد النسبة المئوية للحالات التي تقع مابين أي قيمتين محددتين للمتغير.

حلول التمـــارين

الفصل الأول :

1. a. $\%14 = \frac{127}{903}$

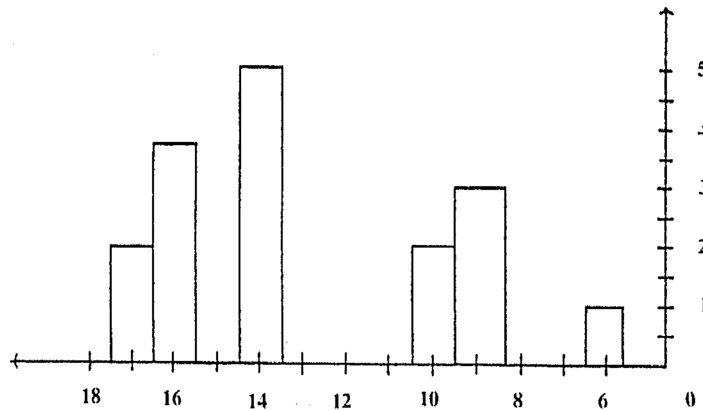
b. المتوسط الحسابي = 154.11 و المدى = 198 ; S = 65.32

2. a. المتوسط الحسابي = 13 ، المتوسط العددي = 14 والمنوال = 14

b. = S = 3.325

c. $Z = 16 - \frac{13}{3.325} = .902$

d.



الفصل الثالث:

$$t = \frac{(10,491 - 10,750)}{\sqrt{\frac{2200}{8} = \frac{(150)^2}{5}}} = \frac{259}{\sqrt{9500}} = 2,657$$

$df = 5 + 8 - 2 = 11$

في حالة المستوى الدلالي الإحصائي يساوي 05.. المنطقة الحرجة قيم t مشابهة لـ 2.201 ـ أو أكـبر من 2.201. واستنتـاجنـا يشير إلى أن هناك اختلاف بين المتوسطات

الحسـابية للمرتبات التي يحصل عليها الرجال والنساء. (فرقان بين المتوسط الحسابي لمرتبات الرجال والمتوسط الحسابي لمرتبات النساء).

الفصل الرابع:

1. $.56 = r$ ، ارتباط ضعيف والمعطيات الحسابية التالية، نتجت عن حل هذه المشكلة الرياضية:

$8 = n$

$29 = \Sigma y$

$121 = \Sigma y^2$

$620 = \Sigma X$

$48,600 = \Sigma X^2$

$2,300 = \Sigma yX$

2. a. $\hat{y} = 776 + ,0955X$

b. $\hat{y} = 3,776 + ,0955(75) = 3,387$.

الفصل الخامس:

1. a.

	أستاذ	أ. مشارك	أ. مساعد	
مستفيدون	53 a	33 b	26 c	112
غير مستفيدين	182 d	97 e	71 f	350
	235	130	97	462

الخليـــة

	a	b	c	d	e	f
fe	56.97	31.52	23.52	178.03	98.48	73.48
fo	53	33	26	182	97	71

$$X^2 = \left(\frac{71^2}{73,48} + \frac{97^2}{98,48} + \frac{182^2}{187,03} + \frac{26^2}{52,23} + \frac{33^2}{31,52} + \frac{53^2}{56}\right) - 462 = ,8$$

$$df = 2$$

النتيجة : لايوجد ارتباط ترتيب أعضاء هيئة التدريس واستخدام المكتبة .

$$b.\ c = \frac{\sqrt{,8}}{462 + ,8} = ,04$$

Abbreviation's List قائمة المختصرات*

Abbreviations of letters مختصرات الحروف

b. Slope	قيمة الانحدار
C. Number of Columns (for df equation)	عدد الأعمدة (في معادلة درجة الحرية)
C. Pearson's Contingency Coefficient	معامل احتمال بيرسون
df, Degrees of Freedom	درجة الحرية
Di, Differences between ranks, in calculating Rs with Interval Scale data	الاختلافات بين ترتيب الأفضليات عند حساب قيمة مع Rs بيانات القياس الحدي
E. The amount of error we will allow for	قيمة الخطأ المسموح به
f, Frequency of occurrence of values or observations	عدد مرات تكرار ظهور القيم
fe., Expected frequency	التكرارات المتوقعة (النظرية)
fo., Observed frequency	التكرارات المشاهدة (الفعلية)
MD., Measures of Distribution	مقاييس التشتت
n.1. Total number of Cases or observations	1. العدد الكلي للحالات البحثية أو المشاهدات
2. Sample Size	2. حجم العينة
N, Total number of elements of the Population	العدد الكلي لمفردات مجتمع الدراسة
ND,-Normal distribution	التوزيع الطبيعي
P., Probability's Value of test of significance	قيمة الاحتمال، لاختبار الدلالة (الاحصائية)
r, (1) Person's Product-moment	(1) معامل ارتباط بيرسون

* [من أعداد المترجمين]

معامل الارتباط الخطي لبيرسون Correlation Coefficient

عدد الصفوف (في معادلة درجة الحرية) (2) number of rows for the df equation

العلاقة بين المتغيرات (3) The relationship between the variables

معامل ارتباط سبيرمان rs, Spearman Correlation Coefficient

تكرار ظهور القيم (مسبوقة أو متبوعة بقيم تختلف عنها) R., Runs

الانحراف المعياري S., Standard deviation

قيمة توزيع T, Tue at t distribution

معامل التغير V. Coefficient of Variation

قيمة كل حالة بحثية أو مشاهدة X, Value of each Case or Observation

المتوسط الحسابي (الوسط الحسابي) للعينة $\bar{X}$

(مفردات العينة) أو المشاهدات (عدد المشاهدات)

مربع كاى (كا) X^2 Chi-Square

الدرجة المعيارية Z Score, Standard Score

(وحدة قياس الانحراف المعياري)

مختصر الرموز — Abbreviation of Symbols

قيمة الـ (رمز إحصائي يسبق مختصرا القيم) Ʃ [Sum of

قيمة التغاير σ , Variance

مربع قيمة التغاير ^{2}σ , Variance Square

القيمة المقدرة للانحراف المعياري σ g, The guessed Value of the Standard Deviation

الانحراف المعياري لـ R σ R, Standard deviation of R.

الخط المعياري للمتوسط الحسابي $\sigma \bar{X}$, Standard error of the mean

المتوسط الحسابي لمجتمع الدراسة $\mu\ \bar{\mu}$ The mean of the Population

المتوسط الحسابي لـ R $\mu_R\ \bar{\mu}_R$ Mean of R.

الكشـــــاف